Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

MASTER THESIS

Chinese Stock Markets: Underperformance and its Determinants

Author: **Bc. Roman Kováč**
Supervisor: **PhDr. Karel Báťa**
Academic Year: **2014/2015**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature and that the thesis has not been used for obtaining another title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 14, 2015

Signature

Acknowledgments

Hereby, I would like to express my gratitude to my supervisor, PhDr. Karel Báťa, for his time and valuable comments. My thanks also belong to my family for support throughout my university studies.

Abstract

Performance of stock markets is determined by three classes of variables: macroeconomic indicators, industry & firm heterogeneity and third country effects. When assessing performance of a stock market index, impact of industry & firm heterogeneity is marginal as it is already embedded in the index through its constituent companies. This paper will therefore focus on the other two. Chinese stock market was selected as an application as their performance compared to other domestic indicators (mainly GDP growth) is considered inferior by many researchers. Using econometric framework for panel data and a Bayesian extension, the paper estimates multiple models of Chinese stock market performance examining individual determinants of it. Subsequently, it predicts development of theoretical prices of two main Chinese stock indices on two time samples until 2013. The paper then demonstrates underperformance of Chinese stock market by comparing the modeled prices to actual prices realized on the market.

JEL Classification	C23, C51, C53, G15, G17
Keywords	underperformance, panel data, fixed effects model, Bayesian Model Averaging

Author's e-mail	roman_kovac@ymail.com
Supervisor's e-mail	karel.bata@seznam.cz

Abstrakt

Výkonnosť akciových trhov je podmienená tromi typmi premenných: makroekonomickými indikátormi, heterogenitou sektoru a firmy a efektami tretích krajín. Pri posudzovaní výkonnosti indexu akciového trhu je dopad heterogenity sektoru a firmy marginálny, pretože tento je už obsiahnutý v indexe skrz firmy, ktoré sú jeho zložkami. Táto práca sa teda sústredí na ostatné dva typy. Čínsky akciový trh bol vybraný pre na ilustráciu vlypvu týchto prvkov, pretože jeho výkonnosť v porovnaní s domácimi ukazovateľmi (najmä rastom HDP) je mnohými výskumníkmi považovaná za relatívne horšiu. Za použitia ekonometrického rámca pre panelové dáta a ich bayesovské rozšírenie, táto práca odhaduje viacero modelov výkonnosti čínskeho akciového trhu hodnotiac jej individuálne determinanty. Model následne predikuje vývoj teoretickej ceny dvoch hlavných čínskych akciových indexov na dvoch časových vzorkách rôznej dĺžky končiace v roku 2013. Práca potom demonštruje nedostatočnú výkonnosť čínskych finančných trhov porovnaním namodelovaných cien s aktuálnymi cenami na trhu.

Klasifikácia JEL	C23, C51, C53, G15, G17
Kľúčové slová	underperformance, panel data, fixed effects model, Bayesian Model Averaging

E-mail autora	roman_kovac@ymail.com
E-mail vedúceho práce	karel.bata@seznam.cz

Contents

List of Tables

List of Figures

Master Thesis Proposal

Author	Bc. Roman Kováč
Supervisor	PhDr. Karel Báťa
Proposed topic	Chinese Stock Markets: Underperformance and its Determinants

Topic characteristics

The size of the Chinese stock market (including stocks listed and traded in Shanghai, Shenzhen and Hong Kong stock exchanges) is the second largest in the world. The underperformance of this market relative to both developed (US) or emerging economies (Brazil, Russia, India and South Africa) has been striking. This is in spite of the fact that Chinese economy (also the second largest in the world) has been the fastest growing global economy for the past three decades.

The poor performance of Chinese stock markets may be attributable to several factors such as undervaluation of Chinese companies or IPO and delisting processes, corporate governance related to self-dealing and information disclosure.

Hypotheses

1. Chinese stock market is underperforming.
2. Undervaluation of Chinese companies is a significant contributor to the underperformance.
3. Internal policies are the significant contributor of the underperformance.

Methodology

The research will be based on a sample of firm-level data from over 90 companies for the period 2008-2014. In order to capture overall Chinese stock market, the sample will be composed from the main indices constituents from three Chinese stock exchanges – Hong Kong Stock Exchange, Shanghai Stock Exchange and Shenzhen Stock Exchange. For purpose of the research, only Chinese companies will be used in the sample. This is especially the case of Hong Kong Stock Exchange with a lot of international listings.

In the first part, underperformance of Chinese stock markets will be examined. The analysis will be based on comparative charts with respect to both internal and external factors. Externally, Chinese stock market performance will be compared to foreign financial markets from both developed and emerging countries. Internal factors used for the comparison will constitute from GDP growth and alternatively performance of other financial market's instruments in China. The analysis will be based on the selected stock markets' indices as a proxy for overall market performance.

Two regression analyses will be used to examine the relationship between the underperformance of the sample and the selected determinants.

The first regression will examine relationship between the sample companies that were undervalued at time of their respective IPOs and their contribution relative to the respective index performance. Undervaluation will be assessed P/V ratios and initial returns of the companies. For the purpose of the regression analysis, the undervaluation will be used as a binary variable.

For the purpose of the second regression, an analysis of internal policies will be carried out. Author finds the companies' corporate strategy as the main aspect of internal policies. To capture the companies' corporate strategy, binary variable examining whether the focus of the sample companies is the profit or size maximization. For this purpose, financial data of the companies will be used as a proxy. Subsequently, regression analysis will be used to examine whether there is a significant relationship between one of the corporate strategies and underperformance of the sample.

Expected contribution

This paper will aim to present determinants of Chinese financial markets underperformance and therefore partly solve the paradox of the fastest growing country (in terms of GDP) having underperforming stock markets. Therefore it will elaborate on underperformance and its main contributors.

From practical point of view, the paper will offer implications under which Chinese stock markets would be performing normally and therefore enable its readers to better target individual companies or indices.

Outline

1. Motivation
2. Studies of Chinese stock market
3. Dataset
4. Methodology
5. Results analysis
6. Conclusion

Core bibliography

CHOW, G.C., 2007. China's Economic Transformation, Blackwell.

FAMA, E. & FRENCH, K., 1989. Business conditions and expected returns on stocks and bonds, Journal of Financial Economics, 25.

QIAO, L. & QIANG, K., 2010. Information-based stock trading, executive incentives, and the principal-agent problem, Management Science, 56(4), 682-698.

QIAO, L. & SIU, A., 2011. Institutions and corporate investment: evidence from an investment-implied return on capital in China, Journal of Financial and Quantitative Analysis.

SHUANG, X., CHANGHUI, Y., HONGJIN, L. & MIN, H., 2009. Appreciation Expectation of RMB and Equilibrium of domestic assets markets, Journal of World Economy.

SOHN, C., TSUI, A., ZHANG, F. & ZHANG, Z., 2012. An Empirical Assessment of A-Share IPO Under-Pricing in China, Seoul Journal of Business, Volume 18, Number 1.

YUNHONG, Y., YAPING, W. & LIANSHENG, W., 2009. Does the stock market affect firm investment in China? – A price informativeness perspective, Journal of Banking and Finance, 33(1): 53-62.

Chapter 1

Introduction

China has been the most outstanding economy during past three decades. It has managed to outperform the world during past three decades. Even though the world’s second biggest economy is now shifting to more moderate pace (yet still high for the rest of the world), it is on verge of overtaking U.S. as a number one (World Economic Outlook, 2015).

While the unprecedented growth of Chinese economy triggered by economic reforms in the 80s has inspired several countries worldwide to follow China, there has been handful of controversial issues emerging throughout the economic boom such as strict government policies (Feng-Cheng et al., 2007).

After inception of Chinese stock markets in the early 90s, there was not too much attention about it especially due to almost complete lockout of foreign investors. However, this has been partly relaxed in the following years and gradually increasing number of foreign investors have turned to China anticipating immense growth opportunities of Chinese markets (Chow, 2007). In reality the expectation were not met as China was a high-inflationary economy with high level of volatility characteristic for emerging countries.

With further economic development and disinflation, Chinese stock market started to finally perform seemingly better. After a period of decent growth, presence of bubbles had knocked Chinese market down again. Although it took very short time to start the recovery, it remains in track-record for the future. Detailed analysis of Chinese stock markets is provided as well as literate review on relationship between GDP growth and stock market performance will be presented in the following two chapters.

This thesis aims to provide retrospective forecast of Chinese stock markets development modelled based on other countries. The dataset and methodology (together with underlying theory) used for this purpose are described in chapters 4 and 5 respectively.

The key part of the thesis is the sixth chapter where the results of both FE model and its BMA extension will be presented. The process of both method will be described step by step together with results obtained by both methods.

The last chapter comments on the results and makes inference about the tested hypotheses that were set in the proposal and specified further in the thesis.

Chapter 2

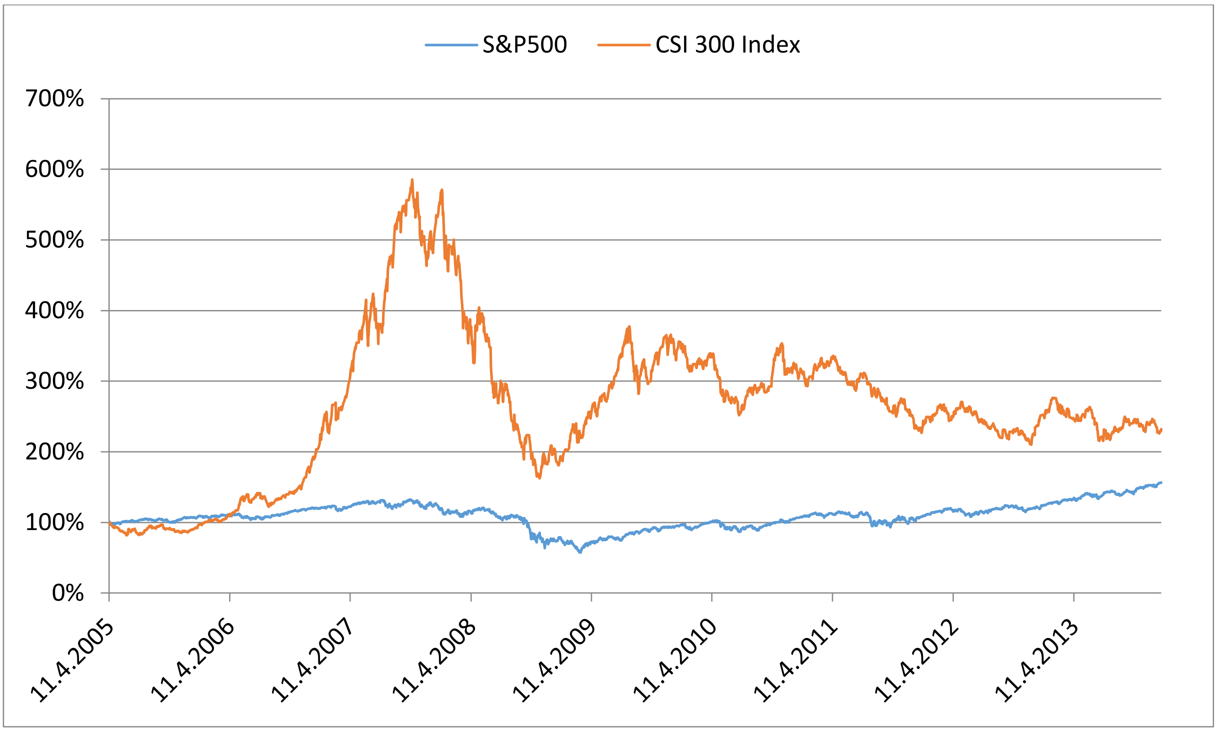
Literature Review

The aim of this chapter is to provide literature review associated with Chinese stock markets underperformance and its main determining factors. Particular stress will be put on GDP growth as an important determinant. Study of the existing literature revealed three sources of determinants that exert influence on performance of stock markets: macroeconomic indicators, industry & firm heterogeneity and third country effects.

This research is working with stock indices as a dependent variable. As an index can be defined as an average firm in an economy, industry and firm heterogeneity is already incorporated in it. Third country effects are included in the models through large sample of countries. Therefore these determinants are reduced to macro-level indicators for the purpose of this research.

One of the main triggers of this thesis is research paper by Allen, Qian, Shan and Zhu (2014). The authors elaborate on Chinese stock market underperformance based on its relationship to Chinese GDP growth compared to other countries. The following paragraphs describe motivation of the paper.

Figure 2.1.: S&P500 vs. CSI 300 Index (2005 – 2013)

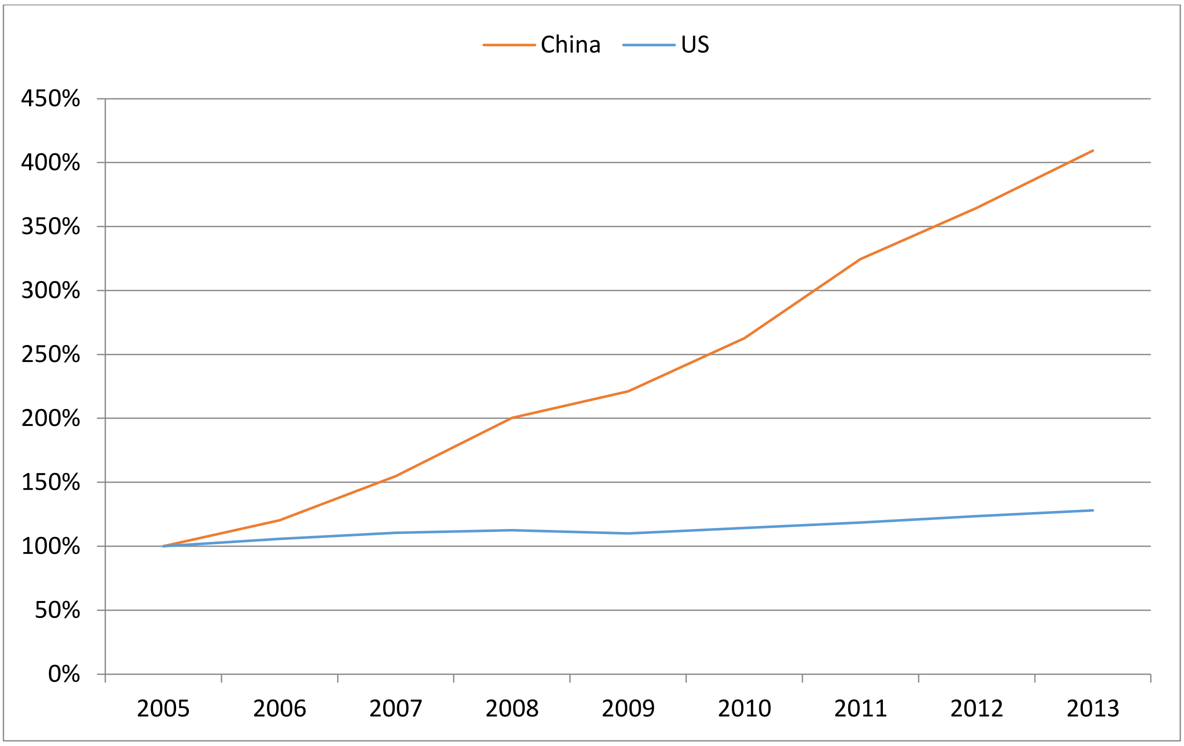


Source: Author's computations

Figure 2.1 depicts the development of the major US and China's stock indices over the period 2005 – 2013. For the purpose of comparability, the values are normalized to April 8 2005 when the CSI 300 Index has been introduced. We can observe the differences between developed US market that shows very little volatility (even during the crisis in 2008) and developing Chinese market that is highly volatile (some of the main reasons being described in previous part of the thesis). Apart from the realized volatility, we can observe convergence of both indices.

In terms of volatility and convergence, the stock indices development seems rather normal. However, compared to GDP growth of both countries throughout he observed period, questions may emerge on whether the two indices should really be converging when the GDP of China grew so significantly more than the GDP of US. Figure x shows this development (normalized to the beginning of observed period).

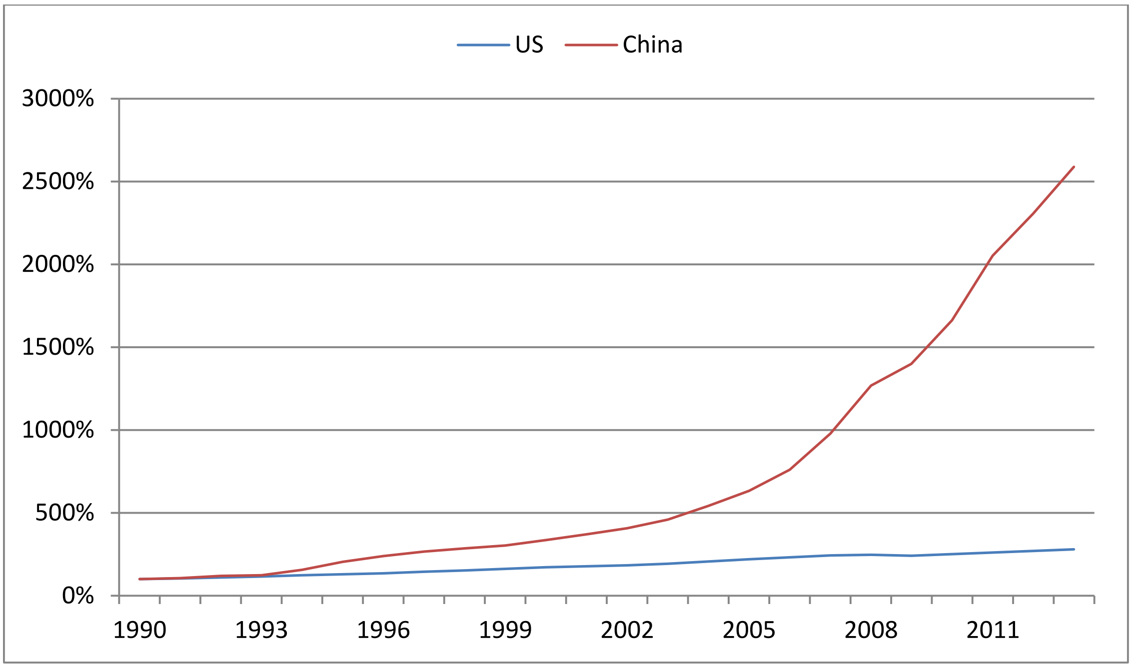
Figure 2.2.: US vs. China GDP growth (2005 – 2013)



Source: Author's computations

When we compare the two countries' GDP over long run, we can observe the same development (with even more striking results). Figure 2.2 shows the GDP development of China and US normalized to 1990 as it was approximately the time when Chinese government economic reforms started to yield effects that have boosted long term growth. It was also the time when Chinese financial markets started to consolidate (resulting in creation of SSE as the first Chinese stock market).

Figure 2.3.: US vs. China GDP growth (1990 – 2013)



Source: Author’s computations

The basic intuition suggests that economic development influence the businesses of individual companies (and therefore their stock prices) through revenues and expenses. From this prospective, changes in economic development can be seen as signals to stock markets development. However, we can observe from the figures above that this was clearly not the case in China.

According to Dimson, Marsh and Staunton (2002), there is no significant evidence of relationship between GDP growth and equity returns in long run. The analysis was performed on shorter panel of historical data of 53 countries but also on 17 countries back to 1990 (Dimson et al., 2002).

However, according to Schroders’ analysis, there was a positive relationship between the two variables during certain part of the business cycle (Wade et al., 2013). This relationship was proven for recovery, expansion and slowdown phases, but not for recession. This justifies need for further research in the topic and partly motivates this paper.

Intuitively, there should be some relationship between GDP growth and stock markets’ development as it directly relates to theory of equity valuation through the concept of discounting that is applied when the stock are valued by the traditional method of discounted cash flows. In this method, expected future cash flows are discounted at the rate that is influenced by interest rates in the respective economy (through risk-free interest rate which is usually approximated by government bonds) (Koller et al, 2005).

Different studies suggest that GDP growth (and eventually other macroeconomic indicators) is not signal to stock price movement but there is rather mutual relationship between the two. It is explained that instead of using GDP growth as a “leading indicator”, it can move in tandem with stock markets or even be lagged (Sandte, 2012). The conclusion of these papers is that economic indicators should not be used to predict future stock prices from various reasons. The first of them is that it is never sure whether the respective macroeconomic indicators run ahead of the stock index. Another reason is that it is simply too demanding to forecast macroeconomic indicators exactly enough to draw conclusions from these predictions. However, this may not cause any flaws to this research as it is using the relationship rather retrospectively.

Based on this theory, we can state that even if there is no or only weak relationship between past GDP growth and stock markets’ returns, there is reason to believe that expected GDP growth have effect on the stock prices (as it is linked through discounting). The paper will also assume that past GDP growth rates can be used as a proxy for its future development (i.e. recent past growth of a country’s GDP implies its growth to foreseeable future).

Chapter 3

Overview of Chinese stock markets

Since initiating market reforms in 1978, China has switched from central planning to market economy and experienced rapid economic and social development. GDP annual growth averaging about 10 percent has lifted more than 500 million people out of poverty (World Bank, 2014). With a population of 1.3 billion, China became the second largest economy and is playing an increasingly important role in the global economy.

The current stock markets in China started their operations in late 1990 with Shanghai Stock Exchange (SSE) followed by Shenzhen Stock Exchange (SZSE) shortly after in 1991. Both exchanges are founded and governed by China Securities Regulatory Commission (CSRC), a joint venture between SSE and SZSE, specialized in creation and management of indexes and index-related services. Even though Hong Kong Stock Exchange has started earlier, it was not part of Chinese stock markets until Hong Kong handover from UK to China in 1997 (Allen et al., 2014).

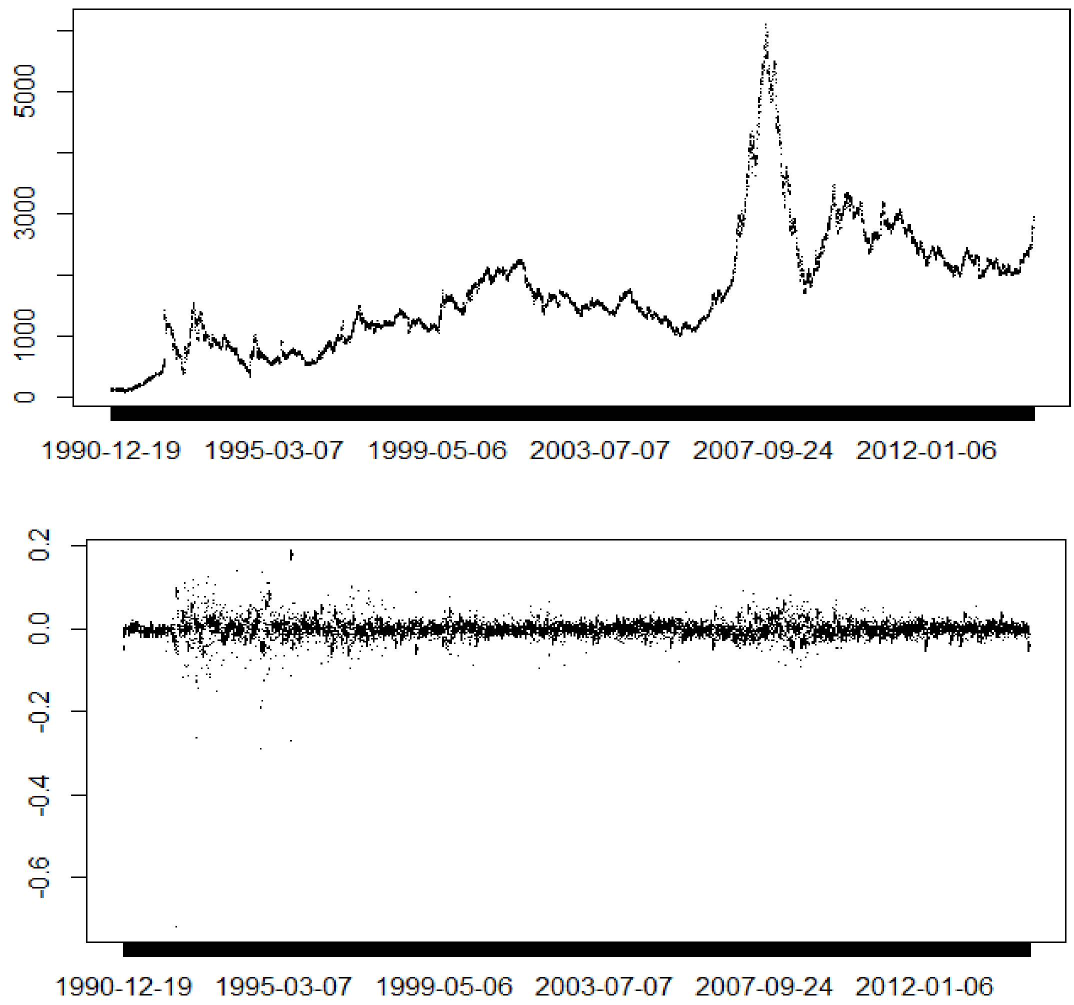
There are two types of shares issued at both SSE and SZSE – “A” shares and “B” shares. “A” shares are quoted in RMB yuan, while “B” shares are quoted in USD. The initial idea was to restrict foreign investors from trading “A” shares. Trading of “B” shares was open to both local and international investors, with certain limitations though. The limitations were relaxed throughout the time resulting in the latest legislative change that aims to merge SSE and HKEX (Asia News, 2014). The change shall facilitate investments into Chinese companies to foreign investors significantly as it was practically impossible to invest into Chinese companies directly at SSE or SZSE due to heavy quotas and strict requirements (Reuters, 2014). However, the initial idea of gradual merging “A” and “B” shares was not implemented yet. There is another type of share being traded at HKEX called “H” shares. These are shares of Chinese companies that are quoted in Hong Kong dollars and are available to foreign investors (China Securities Regulatory Commission, 2014).

Shanghai Stock Exchange

By the end of 2013, there were 997 stocks listed on SSE with a total market capitalization of RMB 15,116.53 billion (2,575.17 billion of shares), decreasing by 4.75% year-to-year, and free-float market capitalization of RMB 13,652.64 billion (2,373.11 billion of shares), up 1.66% from the previous year. The difference between market capitalization and free float is in availability of trading the respective shares. While market capitalization is calculated as a number of shares multiplied by the current stock price of the respective company, free float captures only stocks that are available to public. Therefore we can see that publicly tradable shares amount to 92.15% of total shares and 90.32% of total capital in RMB. A large number of companies across industry, infrastructure and high-tech sectors have raised capital through listing on SSE (SSE, 2014).

As of 31 January 2015, SSE was the 5th largest stock exchange in the world with market capitalization of 3,986 USD billion. That means shift by 23.41% (from 3,230 USD billion) from 7th place as of 30 November 2014 (World Federation of Exchanges, 2015). It also means SSE has overtaken HKEX in the of market cap. This can be attributable to dynamic growth caused by opening to foreign investors.

Figure 3.1.: Shanghai Composite Index - price and return development (1990 – 2014)



Source: Author’s computations

From the introduction of SSE we can observe higher volatility that is characteristic for emerging markets. This is attributable to the fact that China had opened its markets to the public, however, it was still difficult to trade for foreign investors. This initial (predominantly speculative) period was followed by 10 years of relatively stable development after which an immense growth can be seen from the plot. The growth was caused by speculative traders who rushed into SSE and increased its turnover dramatically to become the second largest stock exchange in terms of turnover for a short period of time. During this period, SSE has also reached its all-time high closing price 6,124 points (Yahoo Finance, 2014). After reaching the peak, steep slump followed. That was attributable to Chinese stock bubble (also called

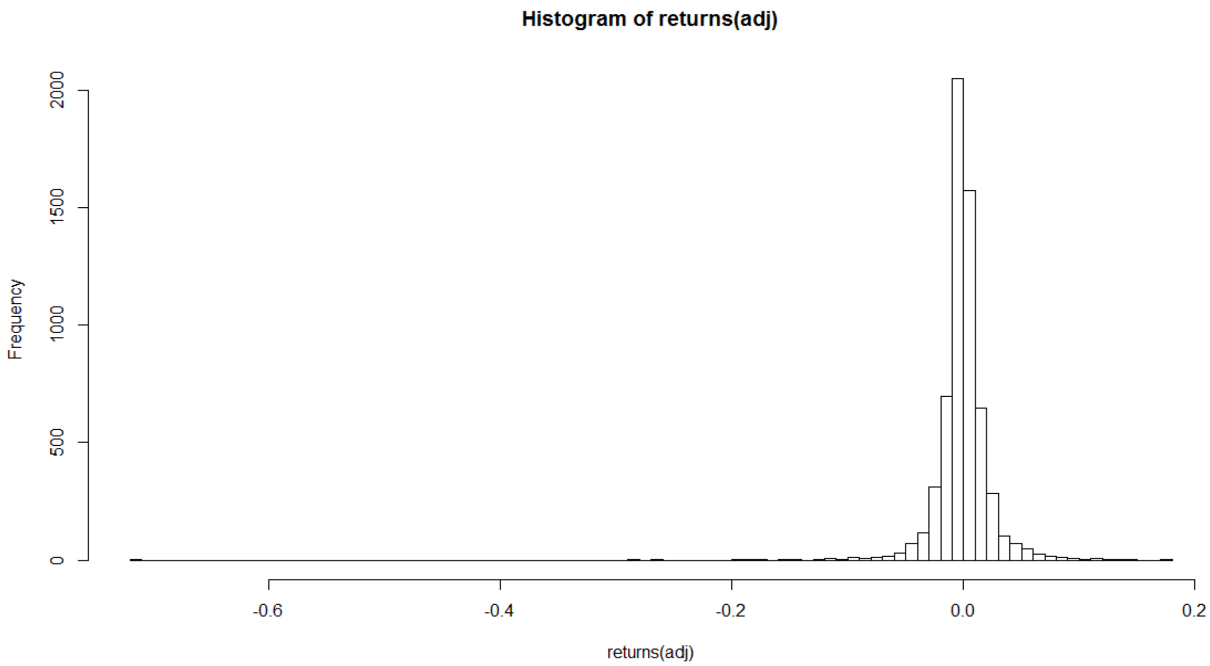
Chinese Correction) that happened in February 2007 due to false expectations of investors who anticipated the Chinese government to raise interest raise in endeavor to tackle inflation. This, however, did not turn into reality and Chinese stock market (and SSE as its main proxy) has fallen sharply. After the initial bounce-back following the bubble burst, the trend of development was steadily declining.

Table 3.1.: Shanghai Composite Index – summary statistics of returns (1990 – 2014)

Min.	-0.7192
1st Qu.	-0.0087
Median	0.0000
Mean	-0.0006
3rd Qu.	0.0074
Max.	0.1791
Skewness	11.6674
Kurtosis	335.8173

Source: Author's computations

Figure 3.2.: Shanghai Composite Index – histogram of daily returns (1990 – 2014)



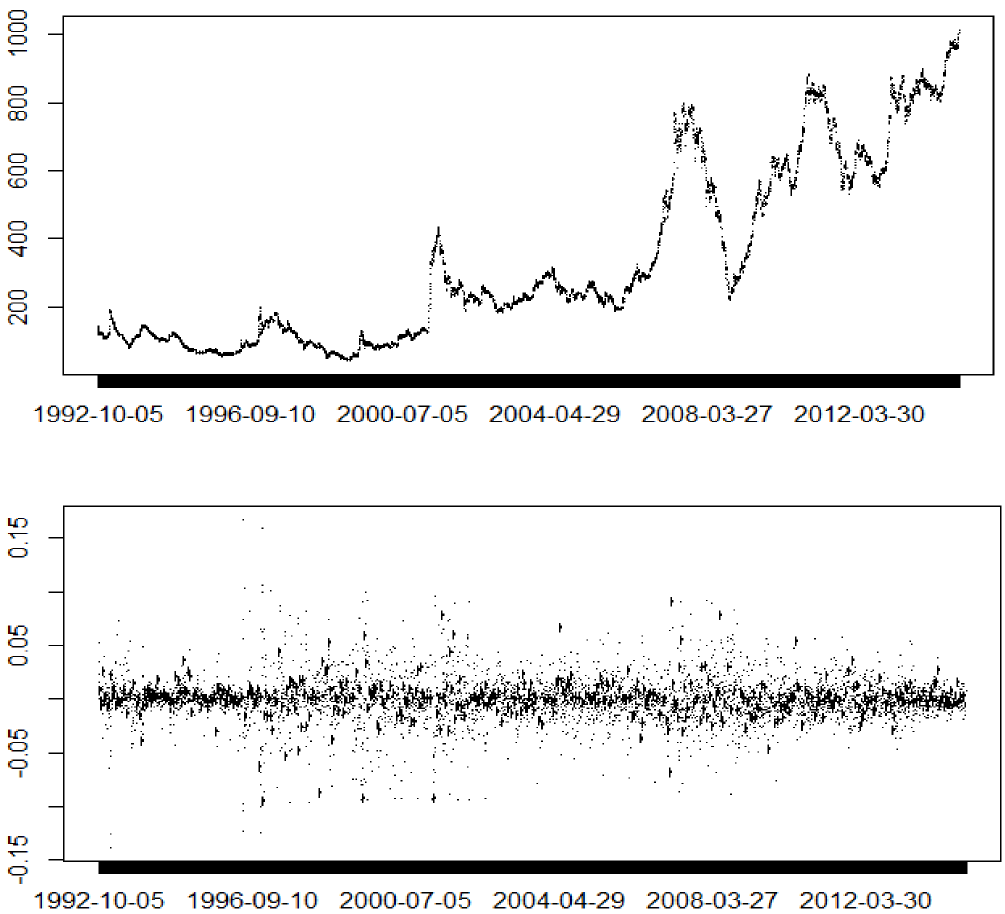
Source: Author's computations

We can observe slightly positively skewed distribution (although the skewness is not very significant) of daily adjusted returns throughout the existence of the stock exchange. This could be partially explained by the presence of extreme negative values, minimum being 71.92% daily loss of the index value.

Shenzhen Stock Exchange

As of 31 January 2015, SSE was the 8th largest stock exchange in the world with market capitalization of 2,285 USD billion and 1,606 listed companies. That means shift by 12.45% (from 2,032 USD billion) from 9th place as of 30 November 2014 (World Federation of Exchanges, 2015). Many of the listed companies are subsidiaries of government-owned enterprises.

Figure 3.3.: Shenzhen Index B - price and returns development (1992 – 2014)



Source: Author’s computations

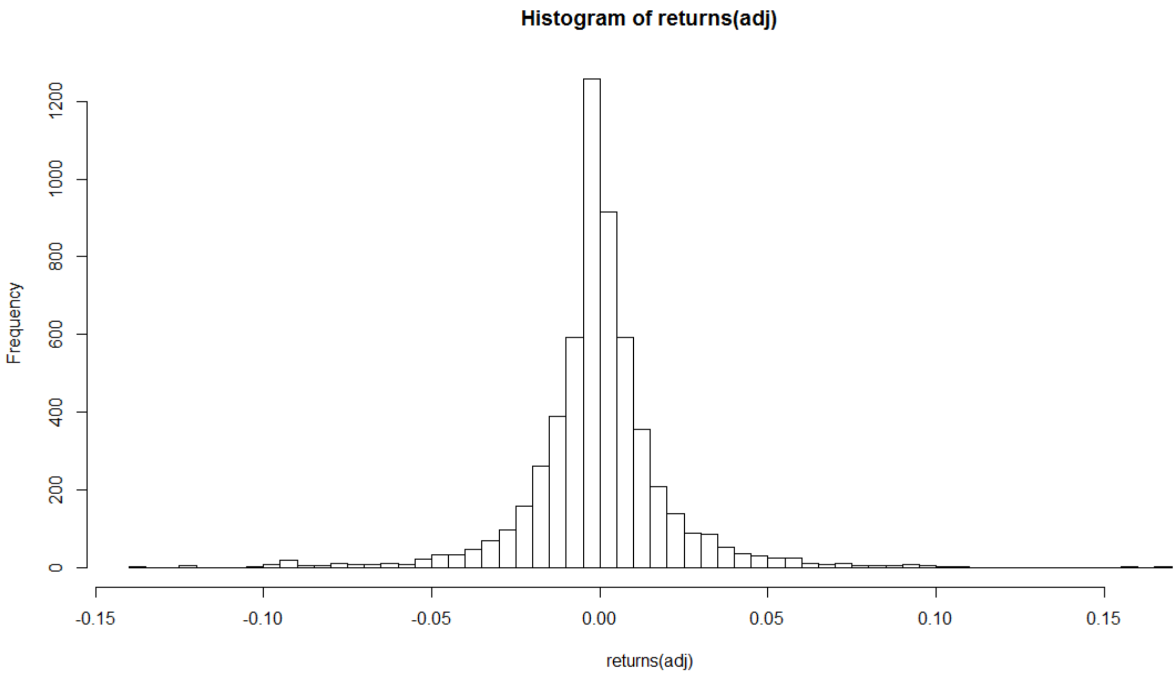
Shenzhen Stock Exchange has seen saddle development throughout the first decade of its operations. The first shock emerged after the peak in 2001 in response to a slump at SSE and as it was mainly related to government-owned companies its effect was more dynamic at SZSE. Later in 2007 the very same effect of Chinese stock bubble like SSE can be observed at SZSE. However, it was followed by a dynamic growth that was shortly interrupted by a steep decline in 2011 which could be attributed by burst of Chinese property bubble (Allan et al., 2014).

Table 3.2.: Shenzhen Index B – summary statistics of returns (1992 – 2014)

Min.	-0.1380
1st Qu.	-0.0078
Median	0.0000
Mean	-0.0003
3rd Qu.	0.0071
Max.	0.1670
Skewness	2.9711
Kurtosis	12.6619

Source: Author's computations

Figure 3.4.: Shenzhen Index B – histogram of daily returns (1992 – 2014)



Source: Author's computations

Similarly to SSE, we can observe slightly positive skewness of the distribution with excess kurtosis and fat tails. However, there is not such a significant outlier as in the case of SSE.

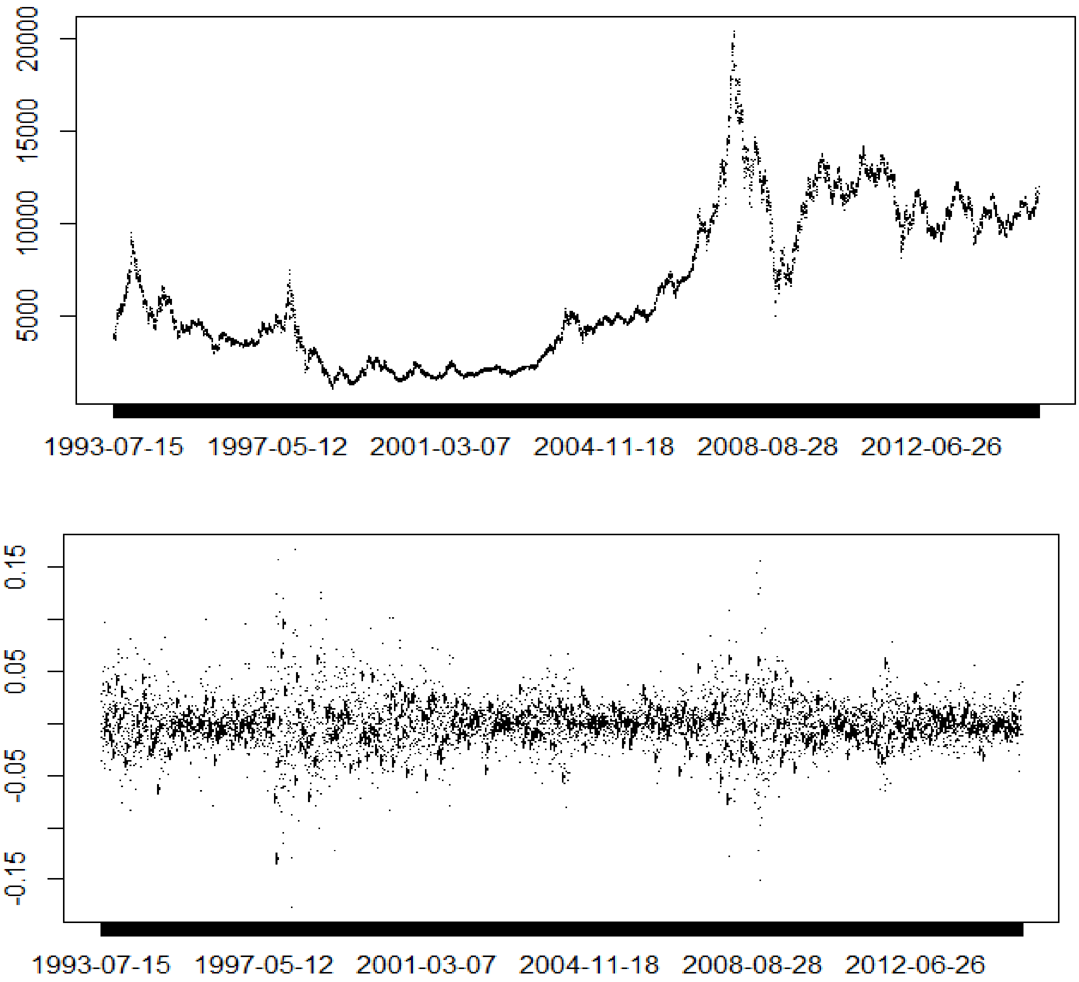
Hong Kong Stock Exchange

Hong Kong Exchange has a long history as it started its operations (of a certain form) under the British rule at the end of 19th century. HKEX in its modern form was founded in 1986 and nowadays it operates the securities and derivatives markets and their related clearing houses in Hong Kong. HKEX is a global leader in base metals trading through ownership of London Metal Exchange (LME) that was acquired in 2012 (LME, 2012).

In terms of regulation, HKEX is accountable to newly established China Exchanges Services Company (CESC) that is a joint-venture of HKEX, SSE and SZSE. However, HKEX has been governed by a stricter regulation similar to those in US or UK throughout its existence and it still keeps following it.

As of 31 January 2015, SSE was the 6th largest stock exchange in the world with market capitalization of 3,325 USD billion and 1,615 listed companies (776 from Mainland China, 737 from Hong Kong and 102 from abroad). That means growth by 1.19% (from 3,286 USD billion) in terms of market capitalization compared to 30 November 2014 (World Federation of Exchanges, 2015).

Figure 3.5.: Hang Seng China Enterprises Index - price and returns development (1993 – 2014)



Source: Author's computations

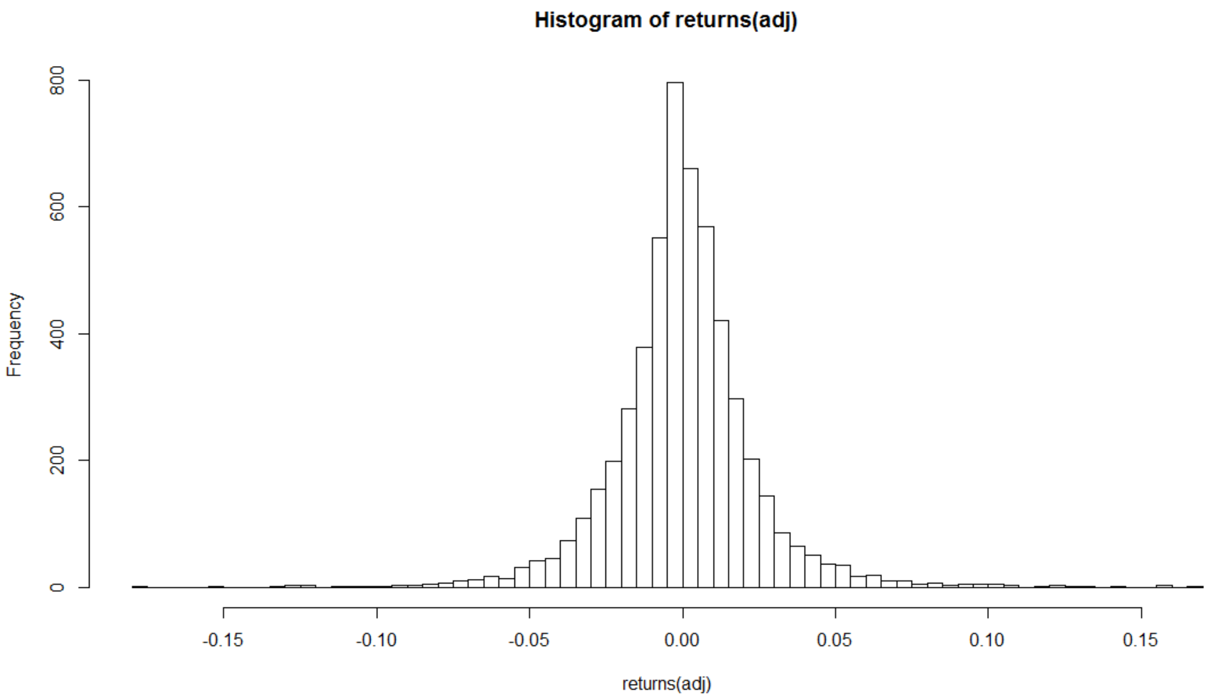
At HKEX we can observe higher volatility in earlier period that was (in spite of couple of peaks) also connected with steady decline. Then, from the late 90s, HKEX has experiences 10 years of accelerating growth that was turned into extreme by speculative investors and (alike SSE) peaked at all-time high of 20,400 points in 2007. The index then fell sharply after the bubble burst and since then it is showing stable (although relatively volatile) development.

Table 3.3.: Hang Seng China Enterprises Index – summary statistics of returns (1993 – 2014)

Min.	-0.1765
1st Qu.	-0.0105
Median	0.0000
Mean	0.0002
3rd Qu.	0.0109
Max.	0.1674
Skewness	2.8967
Kurtosis	14.3629

Source: Author's computations

Figure 3.6.: Hang Seng China Enterprises Index – histogram of daily returns (1993 – 2014)



Source: Author's computations

Unlike the previous two stock markets, HKEX show negative skewness of the daily returns. Other than that, its characteristics are quite similar to SSE and SZSE.

We can observe that all the three stock markets show signs of empirically proven leptokurtic distribution of daily returns. This is characteristic by excess kurtosis compared to normal distribution and fat tails. Excess kurtosis expresses relatively higher probability of returns centered round mean. Fat tails show existence of extreme profits and losses (Tsay, 2002). Leptokurtic distribution is often approximated by normal distribution. However, this approximation can cause significant distortion when examining extreme values (due to the fat tails).

However, at the stock markets, distributions of daily returns usually are slightly negatively skewed (i.e. expected return is slightly above 0). This property is only present at HKEX but not at SSE and SZSE. This is the first indicator of a non-standard behavior of SSE and SZSE.

Chapter 4

Dataset & Variables

Stock market data

As for Chinese stock indices, the most important index, CSI 300 Index, is a capitalization-weighted stock market index designed to replicate the performance of 300 stocks traded in the SSE and SZSE. The index is compiled by the China Securities Index Company, Ltd. It has been calculated since 2005.

Being composed of 300 top “A” shares by market capitalization it also satisfied the condition for high liquidity (most liquid 50% of all “A” shares) and presence of companies across sectors (finance, industry, basic materials, energy, etc.) (CSI Company, 2014). Companies showing great volatility or signs of price manipulation are not admitted to the index. The index contains Chinese companies only, as no foreign companies are allowed to get listed on neither SSE nor SZSE yet (however change is anticipated with upcoming link between SSE and HKEX).

Situation is more difficult when it comes to HKEX. The main HKEX index is Hang Seng Index (HSI). HSI is a capitalization-weighted stock market index in the HKEX. It is used to record and monitor daily changes of the largest companies of the Hong Kong stock market and as the main indicator of the overall market performance in Hong Kong. These companies represent about 70 percent of capitalization of the HKEX. HSI Services Limited is responsible for compiling, publishing and managing the HSI and a range of other HKEX indices.

However, there is a major drawback in using main HSI index for purpose of this analysis. Main HSI index is composed from both Chinese and international companies (or Hong Kong based subsidiaries of foreign companies that are classified as local ones at HKEX). Moreover, the HKEX complies with regulation that is much closer to the West than China. From these reasons, HKEX is not used for predictions and subsequent hypothesis testing.

Information on Chinese stock indices were retrieved from Yahoo Finance. Historical prices of other countries' respective major stock indices were retrieved from Thomson Reuters Eikon. See Annex B for full list of countries and their major stock indices.

GDP growth

Information on annual countries' percentage GDP growth is retrieved from World Bank database (World Bank, 2015). Countries with missing data (e.g. due to war conflicts or changes of territorial structures) are excluded from the sample.

GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. GDP growth is calculated at market prices based on constant local currency (World Bank, 2015).

Corruption

As a proxy for corruption, Corruption Perception Index (CPI) is used. CPI is published annually for 175 countries and territories by Transparency International. The CPI ranks countries and territories based on how corrupt their public sector is perceived to be. It is a composite index (a combination of polls) drawing on corruption-related data collected by a variety of reputable institutions. The index reflects the view of observers from around the world, including experts living and working in the countries and territories evaluated (Transparency International, 2015).

The CPI value range for each evaluated entity is 0 – 10 (10 being the best). In order to capture dynamics, yearly percentage change of the index of a respective entity is used for the purpose of the regression.

Inflation

Information on annual countries' percentage inflation is retrieved from World Bank database (World Bank, 2015).

Inflation is measured by the consumer price index which reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The price change is calculated by the Laspeyres formula

$$P(L) = \frac{E(p_1 * q_0)}{E(p_0 * q_0)}$$

where p0,1 are prices of a consumption bucket at the end of a respective year and q0 is amount at the end of the base year.

Political stability

As a proxy for political stability, this paper uses The Worldwide Governance Indicators (WGI) project by World Bank. The WGI reports aggregate and individual governance indicators for 215 economies over the period 1996 – 2013 (almost perfectly fitting the period analyzed in this paper). The WGI project was run across six dimensions of governance, out of which "Political Stability and Absence of Violence" is relevant for this paper.

Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism (WGI, 2014). The indicator is measured on the scale of approximately -2.5 – 2.5 (higher values being better), although values can occasionally overpass this interval from both sides. To capture dynamics and eliminate possible difficulties in interpretation, this paper uses annual percentage change of the indicator rather than absolute values.

P/E ratio

P/E ratio is calculated based on the stock market data retrieved from Yahoo Finance and Thomson Reuters Eikon. The calculation is done according to the following formula:

$$P/E\ ratio\ =\ price\ of\ index\ /\ earnings\ of\ index$$

P/E ratio is included among explanatory variables as a proxy to determine whether an average company in each economy (expressed by the stock market index) focuses on size or profit maximization. This follows the intuition that higher value of P/E ratio indicates orientation on size while lower P/E ratio should indicate orientation on profit.

Compared to previous explanatory variables, P/E ratio is a stock variable. That means it is represented at one point of time. Due to specific characteristics of P/E ratio, it is not possible to directly convert it into flow variable as it would lack any reasonable economic interpretation and could cause possible flaws in the model. Instead, the paper uses natural logarithmic transformation of absolute value of P/E ratio in the modelling framework. Logarithmic transformation helps to improve model fit and interpretation of the results. Taking absolute value of P/E makes it feasible to compute logarithms and does not negatively affect the model.

Foreign direct investment

Information on countries' foreign direct investment (FDI) is retrieved from World Bank database (World Bank, 2015).

FDI used for this paper are the net inflows (new investments less disinvestments) of investment to acquire a management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. It is calculated as a sum of equity capital, reinvestment of earning, other long-term capital and short-term capital as shown in the balance of payments. Data are in current U.S. dollars as of end of the year.

The paper uses annual percentage change in the net inflow of FDI to capture the dynamics.

Literacy rate

Information on countries' literacy rate is retrieved from World Bank database. Adult literacy rate is calculated as the percentage of population of age 15 and above who can read and write with understanding simple statements of everyday life. Term "literacy" also encompasses "numeracy" which is defined as the ability to make simple arithmetic calculations (World Bank, 2015).

A problem encountered when analyzing literacy rate was that, unlike previous explanatory variables, frequency of measuring literacy in individual countries differs. In practice, it means that while for some countries there are complete annual figures, for others there are big gaps between observations (this is the case in most of the sample). This was solved by simple linear interpolation between the available observations. If there were missing values at the end of the observed period, linear extrapolation was used.

After obtaining balanced panel, annual percentage changes are calculated to capture the dynamics.

Total reserves

Information on countries' total reserves is retrieved from World Bank database. Total reserves comprise holdings of monetary gold, special drawing rights, reserves of IMF members held by the IMF and holdings of foreign exchange under the control of monetary authorities. The gold component of these reserves is valued at year-end (December 31) London prices. Data used are in current U.S dollars (World Bank, 2014).

For the purpose of the regression, annual percentage change is used.

Real interest rate

Information on countries' real interest rates (RIR) is retrieved from World Bank database. Real interest rate is the lending interest rate adjusted for inflation as measured by the GDP deflator (World Bank, 2015).

A major drawback regarding RIR is that some countries do not publish this information and some others stopped publishing it during the examined period. Therefore there is only unbalanced panel available for this explanatory variable. As it is not very convenient from the modelling prospective, author substitutes N/A values with zeros. This substitution, however, incorporates implied assumption of lending interest rate equal to inflation rate. Although this assumption does not strongly deviate from economic intuition (higher inflation tends to increase interest rates), it is necessary to be careful when making inference based on this explanatory variable.

Time period

Time period used for the first regression is 1996 – 2013. This is due to limited availability of data before 1995 (1995 data were often necessary because of dynamics). Out of full sample of 66 countries, there is complete data available for 29 of them from 1996. This sample contains countries from both developed and emerging countries across the continents and therefore it sufficient to make inference.

Alternatively model will be estimated also on shorter timespan 2003 – 2013 with 56 countries' values in the subsample. In spite of original intention to do analysis until the end of 2014, this was not possible due to huge amounts of data yet to be published.

Chapter 5

Methodology

<u>Hypotheses</u>

Hypothesis #1

The underperformance hypothesis is tested by predicting prices of Chinese stock indices using model described further. The modelling is based on longer (1996 – 2013) and shorter (2003 – 2013) time samples of multiples countries. Subsequently, realized index prices will be compared to the predictions to confirm or reject underperformance.

Hypothesis #2

Testing of undervaluation of Chinese companies as a determinant of Chinese stock markets underperformance should have been based on firm-level data of individual indices' constituents. The aim was to sort out undervalued companies based on indicative valuation (e.g. using valuation multiples) and examine their influence on underperformance of individual indices. However, due to unavailability and insufficient reliability of Chinese firm-level data, testing of this hypothesis had to be abandoned.

Hypothesis #3

In order to test corporate strategy of Chinese companies as a determinant of Chinese stock markets underperformance a measurable aspect of corporate strategy had to be chosen. The stress was put on a company choice between size and profit maximization. P/E ratio was selected as an indicator of this strategic decision as lower value of P/E ratio signals higher earnings relative to price, implying ceteris paribus higher focus on profit (and vice versa).

Model framework

Model of stock index performance will be estimated based on the presented dataset (excluding China). As a performance measure (dependent variable), the paper is using stock index returns. The main reason is that unlike asset prices asset returns are stationary and ergodic. That means that a time series of asset returns has a constant mean and variance and its covariance depends only on time span between times t and t+1. Ergodicity means that it is possible to determine characteristics of a process from only one realization (Tsay, 2002). Moreover if we consider average investor and perfectly competitive market (very common assumption in research papers) then the size of an investment does not affect the price change.

Explanatory variables employed in the model are described in the previous section.

After the model is estimated, the resulting coefficients will be applied on Chinese stock index from the beginning of the examined period till the end of 2013. Subsequently the actual price will be compared to the price predicted by the model. Based on the result, the first hypothesis will be accepted or rejected.

In order to capture specific regional or periodical effects, the regression will be run on full sample as well as on subsamples. In terms of geographical characteristics, three subsamples will be used - Asian countries, big countries and BRIC. From the time prospective, the paper will analyze longer period (1996 – 2013) and shorter period (2003 – 2013).

The main determinants will be selected from the explanatory variables based on the estimated model and also evaluated across subsamples.

The paper uses R Studio for data testing and models estimation. Simple calculations are done in MS Excel using outputs from R Studio.

Returns

We can distinguish simple returns and logarithmic returns. Simple net return can be calculated as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

In order to obtain compounded simple net return for multiple periods, it is necessary to multiply underlying gross returns. For k periods, the formula looks as follows:

$$1 + R(k)_t = (1 + R_t) * (1 + R_{t-1}) * \dots * (1 + R_{t-k+1})$$

After subtracting 1 from compounded simple gross return, we get compounded net return. Therefore it is obvious that simple returns are not cumulative, which is not very convenient property as it is often necessary to analyze return over a different time periods. Logarithmic returns are handy way to solve the issue.

Log return (also called continuously compounded return) can be calculated as:

$$r_r = \ln(1 + R_t) = \ln(\frac{P_t}{P_{t-1}})$$

Unlike simple returns, log returns are cumulative, so it is very convenient to work with them. However, log returns are not always the best as log return is rather an approximation (Tsay, 2002).

Another issue in working with returns is dividends as they present sort of an irregularity in asset price. In practice, asset price declines after dividend is paid out. Therefore it is necessary to increase ex-dividend price Pt by the dividend Dt. This problem is eliminated as most of the major financial databases provide asset prices adjusted for dividends paid out (“Adjusted Close Price”).

For the purpose of testing of the first hypothesis, the author is using simple returns as it is more appropriate in the situation when low frequency financial data is analyzed (and therefore bigger changes are expected). Inputs to the regression analysis will be annual GDP growth and annual return of major stock index of the selected countries. In the case of yearly returns simple and log returns differ significantly, therefore approximation by log returns does not come handy. Moreover, only annual data will be used, so there is no need for compound returns (the biggest advantage of using log returns).

It is important to mention a minor drawback of the returns analyzed in the paper. It is desirable to use excess returns for such analyses (i.e. net stock index return less risk free rate of a respective country). However, this was not feasible due to unavailability of reliable information on risk free rates of certain countries and their development in time. Neither it was possible to extract this information from government bonds due to inexistence of appropriate instruments.

The linear panel model

The paper is based on balanced panel data. Panel (or longitudial) data involve repeated observations on a cross-section of individuals over time. Therefore panel data can be formally summarized as:

$$y_{it} \; = \; \alpha_{\,it} + \beta_{it}^{T} \; x_{it} + u_{it}$$

where i = 1,...,n (individual countries), t = 1,...,T (time index) and u is White noise (Woolridge, 2010).

Typically, number of assumptions about the parameters and errors are made when examining panel data. The most common assumption made is homogeneity of α and β, which results in a standard linear modelling framework (pooled regression):

$$y_{it} = \alpha + \beta^T x_{it} + u_{it}$$

The aim of panel data models is, however, to observe differences in behaviour among the cross sections keeping the time dynamics (Woolridge, 2010). Once the homogeneity assumption has been made, the next step is to prove presence of unobserved (individual) effects. The presence of unobserved (individual) effects will be tested using Breusch–Godfrey test. It tests correlation between residuals of the model with idiosyncratic errors. The null hypothesis is serial uncorrelation implying no unobserved effects. Alternative is serial correlation in idiosyncratic errors (Breusch et al., 1980).

The Breusch-Godfrey test performed on the data results in p-value < 2.2*e-16, leading to rejection of the null hypothesis. This confirms there are unobserved effects in the data.

It is therefore necessary to model individual effects. In order to do so, a typical procedure is to devide the error term into two separate components, one of which is specific to the individual effect (to be further specified in the model) and does not change in time. This leads to unobserved effects model:

$$y_{it} = \alpha + \beta^T x_{it} + \mu_i + \varepsilon_{it}$$

The appropriate estimation framework for this model depends on the properties of the two error terms. The idiosyncratic error e is typically assumed to be well-behaved and independent on the other components (White noise). The individual component μ_i, however, can be either correlated on uncorrelated with the regressors (Baltagi, 2001).

In case ui is correlated with the regressors, OLS estimation of β would be inconsistent and it is necessary to use fixed effects model (least squares dummy variable estimator).

In the opposite case when μ_i is uncorrelated with the regressors, OLS estimation is consistent but inefficient. Therefore generalized least squares (GLS) framework is applied. This model is called random effects model. This estimation is based on difference between the two error components which can be done by several methods (Baltagi, 2001).

Model selection

Both of mentioned models dispose with advantages and disadvantages. Fixed effects model gives unbiased estimates of β, however, these are highly variable sample-to-sample. On the other hand, random effects model may produce inconsistent estimates of β with smaller variance (therefore closer to true estimates for whole population) (Greene, 2012). Model selection therefore depends strongly on a researcher's tolerance of inconsistency or high variance.

The Hausman test

The most widely used selection criterion between the models is Hausman specification test. The test is designed to detect violation of random effects model's assumption that the explanatory variables are orthogonal to the unit effects. In case there is no correlation between the explanatory variables and unit effect, the estimators of β obtained by fixed effects model (β_FE) and random effects model (β_RE) should be approximately the same. The Hausman specification test statistic H is then the difference between the two estimators. This can be formalized as follows:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})' \, [Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE})$$

The null hypothesis of the test is that H is ortoghonal, therefore it has chi-squared distribution with the same degrees of freedom as the number of explanatory variables in the model (Clark et al., 2012). If p-value < 0.05, the two models are different enough and null hypothesis is rejected. Therefore random effects model is rejected to be better model than fixed effects model.

P-value > 0.05, however, does not necessarily indicate random effects model is better than fixed effects model. This is due to the fact that in true correlations between covariates and unit effects in not zero. Thus, not rejecting the null does not necessarily imply preference of random effects model over fixed effects model, but it may as well imply insufficient statistical power of the test. In this case Hausman specification test does not facilitate the model selection (Clark et al., 2012).

After application of Hausman specification test on the full sample of the data, we obtained p-value equal to 0.02646, safely indicating rejection of the null hypothesis. Therefore we can state that using fixed effects model is more appropriate given the data.

Fixed effects model

The easiest solution to eliminate heterogeneity from the regression is using first differences. However, this only works on short panels (where t = 1,2) as it removes invariance of estimators in time.

As the paper is working with a long panel and particularly searching for time-invariant estimators, it is necessary to utilize least squares dummy variable estimator obtained by fixed effects model. The main idea of the model is adding to the specification of indicator variables z_j for each unit, such that z_j[i] = 1 if observation i is in unit j, and z_j[i] = 0 otherwise. For better illustration, model for the data used in this paper can include “time dummies” and “country dummies” in panel data to account for unexplained year-to-year or country-to-country variation. In this paper, author will focus on year-to-year variation and therefore utilize “time dummies”.

Fixed effects estimator β_FE equals to within-group (or group means) estimator of β, which can be obtained as follows:

$$\beta^{within} = [S_{xx}^{within}]^{-1} \, S_{xy}^{within}$$

where moment matrices S_{xx}^{within} and S_{xy}^{within} are sum of squares and cross products (variation around group means):

$$S_{xx}^{within} = \sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i.})(x_{it} - \bar{x}_{i.})'$$

$$SS_{xy}^{within} = \sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i.})(y_{it} - \bar{y}_{i.})'$$

Bayesian Model Averaging

One of the objectives of this paper is to find main determinants of a stock index performance. For this purpose, author includes relatively high number of explanatory variables in the model (these are described in detail in the previous section). However, high numbers of explanatory variables may also lead to over-specification of the model, which is not desirable as it negatively influences the model's robustness.

In order to handle this potential drawback, author uses Bayesian Model Averaging (BMA) method. For better understanding, it can be useful to provide brief overview about the method.

BMA addresses an ordinary regression problem of model uncertainty. Suppose following linear model structure:

$$y = \alpha_\gamma + X_\gamma \beta_\gamma + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I)$$

where y is a dependent variable, α_γ a constant, β_γ regression coefficients and ε a normal i.i.d. error term with variance σ^2 (Hoeting et al., 1999).

As already mentioned, a problem arises when there are many potential explanatory variables in a matrix X. It is necessary to determine which variables $X_\gamma \in \{X\}$ should then be included in the model and eventually how important they are. The direct approach to do inference on a single linear model that includes all variables is inefficient or even infeasible with a limited number of observations.

BMA tackles the problem by estimating models for all possible combinations of {X} and constructing a weighted average over all of them. If X contains K potential variables, this means estimating 2^K variable combinations and thus 2^K models. The model weights for this averaging stem from posterior model probabilities (PMP) that arise from Bayes‘ theorem:

$$p(M_\gamma|y,X) = \frac{p(y|M_\gamma,X)p(M_\gamma)}{p(y|X)} = \frac{p(y|M_\gamma,X)p(M_\gamma)}{\sum_{s=1}^{2^K} p(y|M_s,X)p(M_s)}$$

Here, p(y|X) denotes the integrated likelihood which is constant over all models and is thus simply a multiplicative term. Therefore, the PMP $p(M_\gamma|y,X)$ is proportional to the marginal likelihood of the model $p(y|M_\gamma,X)$ (the probability of the data given the model M_γ) times a prior model probability $p(M_\gamma)$ – that is, how probable the researcher thinks model M_γ before looking at the data (Hoeting et al., 1999). Renormalization then leads to the PMPs and thus the model weighted posterior distribution for any statistic θ (e.g. the coefficients β):

$$p(\theta|y,X) = \sum_{\gamma=1}^{2^K} p(\theta|M_\gamma,y,X)p(M_\gamma|X,y)$$

The model prior $p(M\_\gamma)$ has to be elicited by the researcher and should reflect prior beliefs. A popular choice is to set a uniform prior probability for each model $p(M\_\gamma)$ proportional to 1 to represent the lack of prior knowledge (Fernandez et al., 2001). Apart from uniform prior probability, there are also other options depending on strength of prior information on model size. For instance, „random theta“ prior proposed by Ley and Steel (2008) is becoming increasingly popular. The prior suggest a binomial-beta hyperprior on the a priori inclusion probability (Ley et al., 2008). For purpose of this paper, author uses uniform prior probability.

Chapter 6

Results Analysis

Fixed effects model

As described in Methodology section, we will try to prove or reject Chinese stock market underperformance and eventually find its main determinants using FE model.

In the first part, models will be estimated for the period 1996 – 2013 for full country sample and individual subsamples. Subsequently, price development of both SSE and SZSE indices will be modeled using coefficients obtained from the models. The price prediction will be based on the indices' prices from 1995. Final year of the prediction is 2013 and will be compared to the actual index price of the respective index.

The first model is based on full sample containing 29 countries representing developed and developing countries from 4 continents (for full list of the countries see Annex B.1). Results of the regression are shown in the Table 6.1.

Table 6.1.: FE, Full sample model (1996 – 2013)

Residuals				
Min.	1st Qu.	Median	3rd Qu.	Max.
-1.61	-0.201	0.0331	0.191	3.89

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
GDP	0.145444	0.594116	0.244800	0.806710
corr	0.143148	0.256026	0.559100	0.576340
infl	1.374344	0.258974	5.306900	0.001550
polstab	-0.292312	0.358372	-0.815700	0.415090
log(abs(PE))	-0.023155	0.016419	-1.410200	0.159110
FDI	0.001404	0.003764	0.373000	0.709280
literacy	-0.843579	5.105839	-0.165200	0.868840
reserves	0.208925	0.082991	2.517400	0.012140
RIR	0.665512	0.349541	1.904000	0.057510

Total Sum of Squares	77.970
Residual Sum of Squares	72.161
R-Squared	0.074503
Adj. R-Squared	0.069079
F-statistic	4.32913
p-value	0.013065

Source: Author’s computations

The most significant result of this regression is impact of inflation on stock market indices development. This is also the most general result, as we will encounter high impact of inflation with high significance throughout the following models as well. This result is in line with economic theory as the dependent variable is a nominal price of an index, and nominal price as such is influenced by inflation.

The other statistically significant determinant estimated by the model is “reserves”. The interpretation is that a country’s reserves growing by 1 percentage point trigger the stock market index to grow by almost 21 basis points.

The last statistically significant explanatory variable is real interest rate. This is an anticipated result, however, the coefficient itself is unexpectedly high.

Out of other explanatory variables that are not statistically significant, there are two surprising relationships. Based on the model, there is negative influence of political stability and literacy on stock market indices. This can be caused by an emerging market paradox when investors want to invest requiring higher yield go for countries with high volatility. Once these countries start to stabilize (increasing literacy and political stability can be seen as a proxy for stabilization), the volatility and yield decreases and therefore the investors exit.

The second model is based on subsample containing 10 countries from Asia. The aim of running regression based on Asian countries is to capture region-specific effect of each explanatory variable. Countries in this subsample include Hong Kong, India, Indonesia, Japan, Jordan, South Korea, Malaysia, Philippines, Thailand and Turkey. The underlying assumption is that the Asian countries are likely to share common characteristics because of geographical, historical and cultural proximity (relative to other countries included in the full sample). Results of the regression are shown in the Table 6.2.

Table 6.2.: FE, Asian countries subsample model (1996 – 2013)

Residuals

Min.	1st Qu.	Median	3rd Qu.	Max.
-1.73	-0.241	0.0204	0.199	3.75

Coefficients

	Estimate	Std. Error	t-value	Pr(>|t|)
GDP	-0.071252	1.130266	-0.0630	0.949813
corr	0.353278	0.512142	0.6898	0.491311
infl	1.658754	0.387844	4.2769	0.021831
polstab	0.038147	0.818151	0.0466	0.962869
log(abs(PE))	-0.027255	0.032632	-0.8352	0.404838
FDI	0.001167	0.005179	0.2254	0.821969
literacy	-1.082503	7.017513	-0.1543	0.877600
reserves	0.586134	0.188492	3.1096	0.002216
RIR	2.421443	1.003645	2.4126	0.016961

Total Sum of Squares	47.336
Residual Sum of Squares	39.366
R-Squared	0.16837
Adj. R-Squared	0.15060
F-statistic	3.62186
p-value	0.00038759

Source: Author’s computations

This model is in line with the previous one in terms of selection of statistically significant explanatory variables. While inflation and reserves have relatively similar coefficients to the former model, the impact of real interest rate is much stronger.

Another counter-intuitive result is negative impact of GDP growth. However, the standard error is too high to make significant inference based on that.

The third model is based on subsample containing 10 largest economies of the full sample. The aims of running regression based on market size is to capture size-specific effect of each explanatory variable. Countries in this subsample include Brazil, France, Germany, India, Indonesia, Japan, Mexico, Philippines, Turkey and US. Their common characteristic is that all of them has population over 65 million people. Results of the regression are shown in the Table 6.3.

Table 6.3.: FE, Big economies subsample model (1996 – 2013)

Residuals

Min.	1st Qu.	Median	3rd Qu.	Max.
-1.66	-0.243	0.0347	0.198	3.69

Coefficients

	Estimate	Std. Error	t-value	Pr(>|t|)
GDP	-1.231659	1.383321	-0.8904	0.374599
corr	0.210442	0.518327	0.4060	0.685280
infl	1.484975	0.366378	4.0531	0.052879
polstab	0.142452	0.783327	0.1819	0.855926
log(abs(PE))	-0.056651	0.037330	-1.5176	0.131087
FDI	-0.007288	0.020072	-0.3631	0.717016
literacy	-1.863055	7.741753	-0.2407	0.810132
reserves	0.594331	0.220130	2.6999	0.007677
RIR	0.960133	0.597783	1.6062	0.110199

Total Sum of Squares	45.777
Residual Sum of Squares	38.683
R-Squared	0.15496
Adj. R-Squared	0.13860
F-statistic	3.28037
p-value	0.0010764

Source: Author’s computations

Results of the regression are almost identical to the previous models. In this case, impact of real interest rate is not statistically significant. Negative coefficient estimate of GDP growth is even more striking here, yet it still misses statistical significance to draw conclusions.

Theoretical stock index price was predicted by the models for both SSE and SZSE based on the actual prices at 31 December 1995. Table 6.4 summarizes actual and modeled prices of respective indices at the end of 2013.

Table 6.4.: Comparison of predicted and actual prices of SSE and SZSE using FE models (1996 – 2013)

	1995	2013	CAGR
ACTUAL PRICES			
SSE	537.35	2033.08	7.67%
SZSE	60.92	850.11	15.77%
SSE			
PREDICTED PRICE			
world_all		2314.35	8.45%
asia_all		16924.56	21.13%
big_all		553.66	0.31%
SZSE			
PREDICTED PRICE			
world_all		262.38	8.45%
asia_all		1918.76	21.13%
big_all		64.41	0.31%

Source: Author's computations

The prediction results seem very dubious at the first glance. Most importantly, the variance is just too high. Growth, expressed as compound annual growth rate (CAGR), differs significantly accross the models. While the model based on big countries predicts almost no growth, the Asian-based model predicts extremely high growth over the whole eighteen years period.

If we base inference from the models on simple arithmetic mean, which is 9.91%, the conclusion that SSE underperforms and SZSE outperforms would follow. However, given the variance of the models, such inference cannot be considered any significant.

In order to eliminate heterogeneity in time and possibly better results, models will be estimated for shorter period 2003 – 2013 for full country sample and individual subsamples. Similarly, price development of both SSE and SZSE indices will be modeled using coefficients obtained from the models. The price prediction will be based on the indices' prices from 2003. Final year of the prediction is 2013 and will be compared to the actual index price of the respective index.

The first model is based on full sample containing 56 countries representing developed and developing countries from 5 continents (for full list of the countries see annex B.1). Results of the regression are shown in the Table 6.5.

Table 6.5.: FE, Full sample model (2003 – 2013)

Residuals				
Min.	1st Qu.	Median	3rd Qu.	Max.
-0.962	-0.158	0.0175	0.175	2.11

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
GDP	0.229836	0.486819	0.4721	0.6370
corr	-0.072400	0.248543	-0.2913	0.7709
infl	-2.181888	0.541372	-4.0303	0.0428
polstab	-0.366950	0.332518	-1.1035	0.2703
log(abs(PE))	-0.019993	0.014307	-1.3974	0.1628
FDI	0.001930	0.002957	0.6527	0.5142
literacy	0.041026	1.662268	0.0247	0.9803
reserves	0.297213	0.054290	5.4746	0.0449
RIR	0.353707	0.391543	0.9034	0.3667

Total Sum of Squares	79.392
Residual Sum of Squares	71.120
R-Squared	0.104200
Adj. R-Squared	0.093208
F-statistic	7.12164
p-value	0.000388

Source: Author's computations

The model with more countries estimated on shorter time span yield relatively similar results as its longer counterpart. We can observe inflation and reserves being statistically strongly significant while other explanatory variables do not show too much significance.

Looking at the coefficients, we obtain completely different story. While inflation keeps the highest impact on the price (consistent with previous models), it shows inverse relationship in the shorter time span. This may attributable to the financial crisis starting from 2007, which has changed perception of risky investments. We can stipulate that from 2007 on, increasing inflation does not signal economic acceleration but rather higher volatility which can have adverse effect on an investment and therefore can be seen as an incentive to exit.

As in the first part, the second model is based on subsample of Asian countries. The subsample contains 18 countries including Bahrain, HK, India, Indonesia, Israel, Japan, Jordan, Malaysia, Oman, Pakistan, Philippines, Qatar, Saudi Arabia, Singapore, South Korea, Thailand, Turkey and Vietnam. Results of the regression are shown in the Table 6.6.

Table 6.6.: FE, Asian countries subsample model (2003 – 2013)

Residuals					
	Min.	1st Qu.	Median	3rd Qu.	Max.
	-0.856	-0.156	0.0154	0.169	1.08
Coefficients					
		Estimate	Std. Error	t-value	Pr(>|t|)
GDP		-0.163972	0.849347	-0.1931	0.847144
corr		-0.333603	0.385628	-0.8651	0.388200
infl		-2.657887	0.742589	-3.5792	0.000449
polstab		-0.048787	0.566635	-0.0861	0.931489
log(abs(PE))		-0.041384	0.021026	-1.9683	0.050654
FDI		0.002375	0.003657	0.6494	0.516959
literacy		-1.700181	3.088833	-0.5504	0.582744
reserves		0.234635	0.073480	3.1932	0.001675
RIR		-0.739521	0.554595	-1.3334	0.184159
Total Sum of Squares			23.800		
Residual Sum of Squares			19.945		
R-Squared			0.16198		
Adj. R-Squared			0.13989		
F-statistic			3.67257		
p-value			0.00031724		

Source: Author’s computations

Similar to the previous model, we can observe highly significant negative impact of inflation (even higher than in the former) and moderate significant impact of reserves.

Moreover, there is a slightly significant impact of P/E ratio. That means that higher (absolute) value of P/E ratio has negative impact on stock market index. This confirms assumption that focus on profit rather than size (resulting in lower P/E ratio) leads to increasing price. However, both statistical significance and coefficient are too small to reliably prove anything.

Analogically, the third model is estimated on the subsample of big countries. The model contains 15 including Brazil, Egypt, France, Germany, India, Indonesia, Italy, Japan, Mexico, Pakistan, Philippines, Russia, Turkey, US and Vietnam. Results of the regression are shown in Table 6.7.

Table 6.7.: FE, Big economies subsample model (2003 – 2013)

Residuals				
Min.	1st Qu.	Median	3rd Qu.	Max.
-0.816	-0.195	0.0268	0.189	0.989

Coefficients	Estimate	Std. Error	t-value	Pr(>|t|)
GDP	-1.904775	1.190855	-1.5995	0.1119
corr	0.087183	0.431323	0.2021	0.8401
infl	-2.106144	0.981533	-2.1458	0.0336
polstab	0.424618	0.604682	0.7022	0.4837
log(abs(PE))	-0.029764	0.028478	-1.0452	0.2977
FDI	0.019127	0.017995	1.0629	0.2896
literacy	2.942011	2.526990	1.1642	0.2463
reserves	0.550056	0.130770	4.2063	0.0303
RIR	3.972941	0.864063	4.5980	0.0233

Total Sum of Squares	23.233
Residual Sum of Squares	16.161
R-Squared	0.30438
Adj. R-Squared	0.26011
F-statistic	6.85527
p-value	0.001274

Source: Author’s computations

Results of this regression differ in terms of significance, estimating reserves and real interest rate as the most significant variable. Moreover, real interest rate coefficient estimate is extremely high. We can also observe high positive impact of literacy growth (in contrary to the previous models). Its statistical significance, however, does not allow to make strong inference.

The fourth model is estimated based on BRIC countries (Brazil, Russia and India). This model is only available on the shorter timespan as there were a lot of missing data to estimate it for the period 1996 – 2013. BRIC includes four countries that are deemed to be at a similar stale of newly advanced economic development. Even though the true economic proximity is highly questionable, it may be interesting to observe mutual econometric relationships captured by the model.

Table 6.8.: FE, BRIC subsample model (2003 – 2013)

Residuals					
	Min.	1st Qu.	Median	3rd Qu.	Max.
	-0.528	-0.26	-0.0504	0.293	0.532
Coefficients					
	Estimate	Std. Error	t-value	Pr(>|t|)	
GDP	-3.152341	3.161711	-0.9970	0.330097	
corr	0.993194	1.015199	0.9783	0.339051	
infl	3.689583	2.759892	1.3369	0.195572	
polstab	0.349129	1.543114	0.2262	0.823195	
log(abs(PE))	-0.016452	0.070764	-0.2325	0.818410	
FDI	-0.233202	0.158409	-1.4722	0.155808	
literacy	-6.503784	7.671217	-0.8478	0.4061	
reserves	1.110953	0.299070	3.7147	0.001283	
RIR	2.563566	1.300783	1.9708	0.062072	
Total Sum of Squares		6.8739			
Residual Sum of Squares		2.8839			
R-Squared		0.58046			
Adj. R-Squared		0.36938			
F-statistic		3.22833			
p-value		0.012882			

Source: Author’s computations

This regression does not yield very significant results. Only reserves and (partly) real interest rate are estimates to be statistically significant. This may be due to the fact that only three countries are included and therefore sample is too small to make inference.

Stock index price was predicted by the models for both SSE and SZSE based on the actual prices at 31 December 2003. Table 6.9 summarizes actual and modeled prices of respective indices at the end of 2013.

Table 6.9.: Comparison of predicted and actual prices of SSE and SZSE using FE models (2003 – 2013)

	2002	2013	CAGR
ACTUAL PRICES			
SSE	1499.81	2033.08	3.09%
SZSE	209.3	850.11	15.05%
SSE PREDICTED PRICE			
world_BMA		1828.17	2.00%
asia_BMA		450.54	-11.33%
big_BMA		557.44	-9.42%
BRIC_BMA		1633.25	0.86%
SZSE PREDICTED PRICE			
world_BMA		255.12	2.00%
asia_BMA		62.87	-11.33%
big_BMA		77.79	-9.42%
BRIC_BMA		1633.25	0.86%

Source: Author's computations

The prediction results seem somewhat more consistent than the longer time sample, however, they are still dubious. First of all, the variance is still very high. Simple arithmetic mean of the resulting CAGR is -4.47% which is also very suspicious due to negative skewness of stock markets that is empirically observed. Moreover, fact that models based on Asian and big countries yield significant negative growth lacks economic intuition.

The models estimated on shorter time sample seem more accurate in terms of statistical measure (R-squared in particular). Nevertheless, their results do not seem very convincing. If we still decide to draw conclusion based on them, it follow that both SSE and SZSE outperformed as they were growing at positive pace.

Although all the models have led to some inference, it was very weak. The author stipulates this is due to over-specification of the models. Including as many as nine variables in a model where huge heterogeneity among the countries is expected lead to flaws to the model causing inplausible results.

Bayesian Model Averaging

The FE models were estimated on a relatively big sample of countries. Due to significant heterogeneity among the countries, it is reasonable to suspect that a model built upon 9 explanatory variables may be over-specified and therefore not accurate to predict Chinese stock markets performance.

In order to get rid of the over-specification issue and to find the best possible (linear) model, BMA is applied. Analogically to previous subsection, the paper firstly uses full country sample for period 1996 – 2013. In the next step BMA is applied on the data sample. See Table 6.10 for summary statistics of the model.

Table 6.10.: BMA: PIP, Full sample model (1996 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
infl	1.000000	1.512380	0.185389	1.000000
reserves	0.497479	1.299199	0.110778	1.000000
RIR	0.162562	0.721884	0.143298	1.000000
PE	0.076980	0.017215	0.000120	1.000000
polstab	0.054312	-0.183140	0.096944	0.000000
corr	0.050653	0.383788	0.063974	1.000000
FDI	0.043196	0.025714	0.000765	1.000000
Literacy	0.042102	0.048430	0.668947	0.450833
GDP	0.041948	-0.120840	0.105323	0.056805

Source: Author's computations

The first column shows posterior inclusion probabilities (PIP) which quantifies importance of the variables in explaining the data. The PIP value expresses percentage probability of including a particular explanatory variable in the best model. It is calculated as a sum of posterior model probabilities (PMP) where a variable was included. From this criterion, we can observe that inflation variable is included in all posterior models. Apart from inflation, growth of reserves seems to have some explanatory power.

The second column shows posterior mean that is averaged coefficient for each variable across all models (including models where the variable was not included). Looking at this criterion, we come to a similar result of inflation being the most important explanatory variable, just followed by the reserves. This is not surprising results, as there is a direct relationship between PIP and posterior mean value of a coefficient.

Posterior standard deviation is calculated in a similar manner as posterior mean thus averaging posterior standard deviation of each variable across all models. The values of standard errors are slightly higher but not high enough to reliably deny explanatory power of the variables.

Finally, the fourth column shows posterior probability of a positive coefficient expected value conditional on inclusion which in simple terms means certainty of a coefficient sign. In this aspect, six out of nine variables are certain in terms of sign.

Other basic information are shown in the following table.

Table 6.11.: BMA: Other statistics, Full sample mode (1996 – 2013)

Mean no. regressors	1.9692
Draws	512
Burnins	0
Time	0.3120179 secs
No. models visited	512
Modelspace 2^K	512
% visited	100
% Topmodels	98
Corr PMP	NA
No. Obs.	522
Model Prior	uniform / 4.5
g-Prior	UIP

Source: Author's computations

The most important information provided by the Table 6.11 is mean number of regressors. In fact, this determines posterior model size of the best model.

In order to assess which models perform the best, we can sort models according to PMP. Individual models are characterized by binary representation of each explanatory variable. The best five models are shown in the Table 6.12.

Table 6.12.: BMA: Best models based on PMP, Full sample model (1996 – 2013)

	1	2	3	4	5
GDP	0	0	0	0	0
corr	0	0	0	0	0
infl	1	1	1	1	1
polstab	0	0	0	0	0
PE	0	0	0	0	1
FDI	0	0	0	0	0
literacy	0	0	0	0	0
reserves	0	1	1	0	1
RIR	0	0	1	1	0
PMP	0.307885	0.301565	0.059890	0.058461	0.025530

Source: Author’s computations

It is clear that the best model only contains inflation, however, the second best has only slightly lower PMP. It may thus be desirable to also include reserves in the prediction model. Mean number of regressors supports this conclusion, therefore the final model will be using the two explanatory variables.

See the Table 6.13 for the actual estimated posterior coefficients of the selected model.

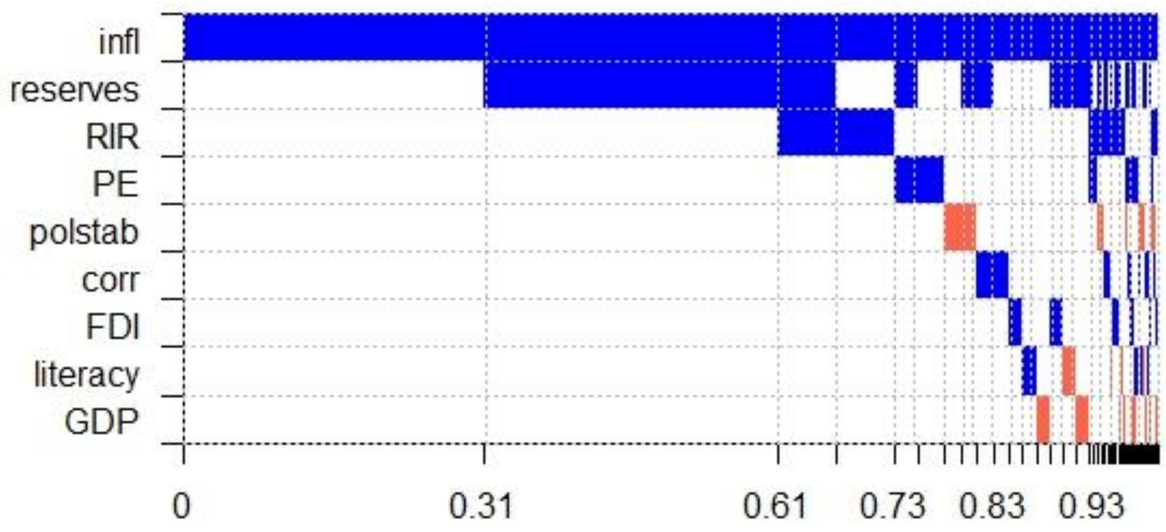
Table 6.13.: BMA: Estimated posterior coefficients, Full sample model (1996 – 2013)

	2
infl	1.500533
reserves	0.192735

Source: Author’s computations

For more complex image about the model probabilities, see Figure 6.1.

Figure 6.1.: BMA: Cumulative Model probabilities, Full sample model (1996 – 2013)

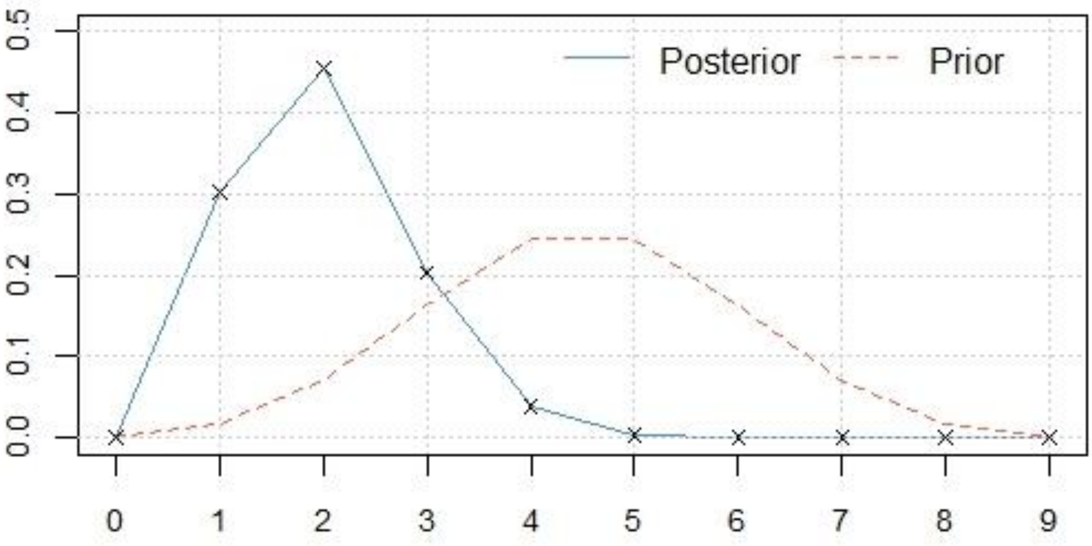


Source: Author's computations

The table depicts inclusion of each variable in the model based on best 121 models. Blue color means positive coefficient, red color mean negative one and white color means zero coefficient (explanatory variable not included in the particular model). The best models can be distinguished by the most mass (range on the x axis).

It is also interesting to observe prior vs. posterior model size distribution. This is depicted in Figure 6.2.

Figure 6.2.: BMA: Prior vs. Posterior model size distribution, Full sample model (1996 – 2013)



Source: Author's computations

As mentioned in the methodology section, the paper is using uniform distribution as a prior for each variable's inclusion in the final model. That means each variable has 50% chance to be or not included in the best model. Therefore it is not uniform distribution of all possible model sizes (in that case the prior model size distribution would be a straight line as a usual uniform distribution suggests).

Cumulative model probabilities and posterior model size distribution graphs for the further estimated models can be found in Annex A.

The same procedure is applied on individual subsamples. BMA output for Asia subsample (1996 – 2013 period) is shown in the following table.

Table 6.14.: BMA: PIP, Asian countries subsample model (1996 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
infl	0.999999	1.731207	0.30657	1
reserves	0.925223	1.959166	0.228835	1
RIR	0.647719	1.485903	1.314724	1
corr	0.105151	0.634467	0.204573	0.999999
PE	0.081549	0.0162	0.000164	1
GDP	0.072551	0.259531	0.273723	0.996224
polstab	0.071904	0.174614	0.209572	0.947892
FDI	0.071622	0.062183	0.001344	0.999999
literacy	0.070739	-0.8394	1.283442	0.053943

Source: Author's computations

We can see that in terms of PIP, inflation and reserves are again the best explanatory variables, real interest rates having relatively high PIP too. Moreover, mean number of regressors given by the model suggest that using three explanatory variables will yield the best results. However, due to quite high posterior standard deviation, any model containing RIR as a regressor is likely to be biased.

Table 6.15.: BMA: Best models based on PMP, Asian countries subsample model (1996 – 2013)

	1	2	3	4	5
GDP	0	0	0	0	0
corr	0	0	1	0	0
infl	1	1	1	1	1
polstab	0	0	0	0	0
PE	0	0	0	1	0
FDI	0	0	0	0	0
literacy	0	0	0	0	0
reserves	1	1	1	1	0
RIR	1	0	1	1	1
PMP	0.369989	0.193421	0.042264	0.032231	0.028953

Source: Author's computations

In spite of the mean number of regressors resulting from BMA, the author decided to only include two explanatory variables in the model. This was due to the combination of high posterior standard error and high coefficient of the third regressor (RIR) that would cause huge distortion had it been included in the model. The coefficient of the remaining two regressors are shown in Table 6.16.

Table 6.16.: BMA: Estimated posterior coefficients, Asian countries subsample model (1996 – 2013)

	2
infl	1.543764
reserves	0.575055

Source: Author's computations

BMA PIPs for subsample of big countries (1996 – 2013 period) is shown in Table 6.17.

Table 6.17.: BMA: PIP, Big economies subsample model (1996 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
infl	0.999999	1.582684	0.262053	1
reserves	0.471724	0.213066	0.264867	1
RIR	0.131879	0.042826	0.147963	1
PE	0.107102	0.000066	0.000281	1
GDP	0.081959	-0.053227	0.369383	0
corr	0.078615	0.019432	0.149536	0.999999
FDI	0.075921	-0.000654	0.005806	0
polstab	0.070015	-0.00623	0.196682	0.064897
literacy	0.069661	-0.017521	1.342138	0.503665

Source: Author's computations

We obtain very similar results of approximately two regressors included in the model, inflation and reserves showing the highest PIP values.

Table 6.18.: BMA: Best models based on PM, Big economies subsample model (1996 – 2013)

	1	2	3	4	5
GDP	0	0	0	0	0
corr	0	0	0	0	0
infl	1	1	1	1	1
polstab	0	0	0	0	0
PE	0	0	0	0	1
FDI	0	0	0	0	0
literacy	0	0	0	0	0
reserves	0	1	0	1	0
RIR	0	0	1	1	0
PMP	0.279224	0.244798	0.040212	0.039798	0.036458

Source: Author's computations

Again, we came to a similar conclusion of including inflation and reserves in the model because of high PMP together with better specification than a single-regressor model. Resulting posterior coefficients are given in the following table.

Table 6.19.: BMA: Estimated posterior coefficients, Big economies subsample model (1996 – 2013)

	2
infl	1.555227
reserves	0.44926

Source: Author's computations

Using the estimated models theoretical stock index price was modeled for both SSE and SZSE based on the actual prices at 31 December 1995. Table 6.20 summarizes actual and modeled prices of respective indices at the end of 2013.

Table 6.20.: Comparison of predicted and actual prices of SSE and SZSE using BMA (1996 – 2013)

	1995	2013	CAGR
ACTUAL PRICES			
SSE	537.35	2033.08	7.67%
SZSE	60.92	850.11	15.77%
SSE PREDICTED PRICE			
world_BMA		2181.87	8.10%
asia_BMA		9622.56	17.39%
big_BMA		6059.32	14.41%
SZSE PREDICTED PRICE			
world_BMA		247.36	8.10%
asia_BMA		1090.92	17.39%
big_BMA		686.95	14.41%

Source: Author's computations

We can see that SSE underperformed compared to the index price prediction by all the models in long run. Full sample model suggests moderate underperformance, the other two show the theoretical price should have been significantly higher. While SSE actual compound annual growth rate (CAGR) was 7.67%, models estimate it should have been 13.3% on average. Even if we do not consider arithmetic average of the three models accurate measure, each value is higher than the actual price or growth respectively.

Looking at SZSE, we can see a different story. Actual CAGR of its returns was more than twice higher than SSE, so if the models are correct, SZSE has outperformed the predicted development. Nevertheless, SZSE slight outperformance does not compensate significant underperformance of SSE. Moreover, SSE was much bigger throughout whole analyzed period (more significantly at the beginning), so the theoretical weighted underperformance would be even more striking.

Now the models will be estimated in the very same manner on the shorter time sample. Below, we can see BMA application summary on the full country sample for period 2003 – 2013.

Table 6.21.: BMA: PIP, Full sample model (2003 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
reserves	0.999999	1.142535	0.050631	1
RIR	0.582674	1.206028	0.32081	1
GDP	0.103625	0.778815	0.209396	1
polstab	0.054301	-0.19742	0.096205	0
literacy	0.045498	0.502462	0.347292	1
FDI	0.043137	0.038717	0.000644	0.999999
infl	0.04094	-0.08777	0.069464	0.497408
PE	0.040618	-0.00462	0.000008	0
corr	0.039053	-0.05997	0.047454	0

Source: Author's computations

Mean number of regressors given by the model remains the same compared to the previous models. However, inflation is no longer the top explanatory variable in terms of PIP.

Table 6.22.: BMA: Best models based on PMP, Full sample model (2003 – 2013)

	1	2	3	4	5
GDP	0	0	1	1	0
corr	0	0	0	0	0
infl	0	0	0	0	0
polstab	0	0	0	0	1
PE	0	0	0	0	0
FDI	0	0	0	0	0
literacy	0	0	0	0	0
reserves	1	1	1	1	1
RIR	1	0	1	0	1
PMP	0.396688	0.288182	0.048858	0.029966	0.02233

Source: Author's computations

Based on PMP, the author selected model consisting of reserves and RIR. Estimated coefficients of the variables are shown in the following table.

Table 6.23.: BMA: Estimated posterior coefficients, Full sample model (2003- 2013)

	1
reserves	0.308303
RIR	0.56089

Source: Author's computations

Results for subsample composed of Asian countries look rather similar to longer time sample models as shown in Table 6.24.

Table 6.24.: BMA: PIP, Asian countries subsample model (2003 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
reserves	0.977428	0.231791	0.073756	1
infl	0.219345	-0.162368	0.37163	0
RIR	0.137321	0.068633	0.231108	1
corr	0.080749	-0.018875	0.121786	0
FDI	0.074055	0.000119	0.001049	0.975972
PE	0.073383	-0.000001	0.000011	0
literacy	0.07055	0.060949	0.737117	1
GDP	0.068249	-0.010293	0.177883	0.022682
polstab	0.068022	-0.008445	0.143791	0.022165

Source: Author's computations

Mean number of regressors given by the model is two and variables with the highest PIP are reserves and inflation. This results is in line with the models estimated on the longer time sample, even though inflation variable seems to be less important in the short run.

Table 6.25.: BMA: Best models based on PMP, Asian countries subsample model (2003 – 2013)

	1	2	3	4	5
GDP	0	0	0	0	0
corr	0	0	0	1	0
infl	0	1	0	0	0
polstab	0	0	0	0	0
PE	0	0	0	0	0
FDI	0	0	0	0	1
literacy	0	0	0	0	0
reserves	1	1	1	1	1
RIR	0	0	1	0	0
PMP	0.411077	0.126167	0.075506	0.037583	0.032662

Source: Author's computations

Table 6.26.: BMA: Estimated posterior coefficients, Asian countries subsample model (2003 – 2013)

	2
infl	-0.745098
reserves	0.230778

Source: Author's computations

Model containing the two variables with highest PIP was finally selected. However, there is a remarkable difference in comparison with the long-run models. It is a rather unintuitive result of negative coefficient of inflation. Nevertheless it is not possible to draw sensible conclusion about the sign as conditional posterior sign value equals zero.

Short-run model of big economies' subsample tells another story. BMA is summarized in the table below.

Table 6.27.: BMA: PIP, Big economies subsample model (2003-2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
reserves	0.992662	0.516651	0.139159	1
literacy	0.295375	1.326095	2.442735	1
RIR	0.139492	0.053298	0.177446	1
FDI	0.102405	0.001678	0.007817	1
GDP	0.080413	-0.032271	0.294998	0.00756
PE	0.080142	0.00002	0.000159	1
infl	0.078258	0.019196	0.195809	0.979805
corr	0.07614	0.011772	0.130809	1
polstab	0.072184	-0.002513	0.173368	0.080584

Source: Author's computations

Even though the mean number of regressors remains the same, for the first time we can see higher PIP of literacy. Reserves seem to be the strongest driver of stock market performance in short run.

Table 6.28.: BMA: Best models based on PMP, Big economies subsample model (2003 – 2013)

	1	2	3	4	5
GDP	0	0	0	0	0
corr	0	0	0	0	0
infl	0	0	0	0	0
polstab	0	0	0	0	0
PE	0	0	0	0	1
FDI	0	0	0	1	0
literacy	0	1	0	0	0
reserves	1	1	1	1	1
RIR	0	0	1	0	0
PMP	0.362687	0.151053	0.05699	0.042332	0.031918

Source: Author's computations

Table 6.29.: BMA: Estimated posterior coefficients, Big economies subsample model (2003 – 2013)

	2
literacy	4.45475
reserves	0.527836

Source: Author's computations

The selected model contains two variable again, both having positive coefficients with absolute certainty. The coefficient of literacy may seem rather high, but it is necessary to understand its high value is offset by very low changes in literacy.

The last short-run subsample consists of BRIC countries and model is estimated on the sample of three countries (Brazil, Russia and India) as it is to be applied on China afterwards.

Table 6.30.: BMA: PIP, BRIC subsample model (2003 – 2013)

	PIP	Post Mean	Post SD	Cond.Pos.Sign
reserves	0.92545	0.850666	0.380475	1
FDI	0.621572	-0.204258	0.200744	0
GDP	0.438741	-1.73037	2.495759	0
infl	0.302133	0.909582	1.872273	1
corr	0.260054	0.291245	0.705442	1
PE	0.168354	-0.000332	0.001714	0.000844
RIR	0.162848	0.000334	0.196142	0.386118
literacy	0.159504	0.063965	2.706556	0.443936
polstab	0.150501	-0.022116	0.609079	0.358309

Source: Author's computations

We can observe the highest mean number of regressors out of all models. This model is likely to be subject of small sample bias as it is apparently difficult to exactly set PIP of each variable (relatively high PIP values of all variables).

Table 6.31.: BMA: Best models based on PMP, BRIC subsample model (2003 – 2013)

	1	2	3	4	5
GDP	0	1	1	0	0
corr	0	0	0	0	1
infl	0	0	0	1	0
polstab	0	0	0	0	0
PE	0	0	0	0	0
FDI	1	0	1	1	1
literacy	0	0	0	0	0
reserves	1	1	1	1	1
RIR	0	0	0	0	0
PMP	0.097853	0.059926	0.057644	0.036211	0.034838

Source: Author's computations

Table 6.32.: BMA: Estimated posterior coefficients, BRIC subsample model (2003 – 2013)

	1
FDI	-0.358275
reserves	0.956776

Source: Author's computations

We can see it is complicated to state which model is the best as PMP values are relatively similar. Therefore the author chose to pick two variables with high PIP and relatively low posterior standard deviation. Their estimated posterior coefficient are shown in Table 6.31.

Using the estimated models theoretical stock index price was modeled for both SSE and SZSE based on the actual prices at 31 December 2002. Table 6.32 shows actual and modeled prices of respective indices at the end of 2013.

Table 6.33.: Comparison of predicted and actual prices of SSE and SZSE using BMA (2003 – 2013)

	2002	2013	CAGR
ACTUAL PRICES			
SSE	1499.81	2033.08	3.09%
SZSE	209.3	850.11	15.05%
SSE PREDICTED PRICE			
world_BMA		3876.3	9.96%
asia_BMA		2306.61	4.40%
big_BMA		7283.18	17.12%
BRIC_BMA		8370.91	18.76%
SZSE PREDICTED PRICE			
world_BMA		540.94	9.96%
asia_BMA		321.89	4.40%
big_BMA		1016.38	17.12%
BRIC_BMA		1168.18	18.76%

Source: Author's computations

In terms of measuring underperformance as a difference between predicted and actual price or CAGR, SSE underperforms even more significantly on shorter time sample. Simple arithmetic mean of CAGR predicted by the models is 12.56% which is more than four times higher than the actual CAGR.

SZSE performed slightly better than predicted by the models. However, similar to the longer time sample, this is easily offset by larger underperformance of SSE.

Overall, the models obtained from BMA seem more legit as the predicted values are much more realistic and main determinants are consistent across the models.

Chapter 7

Conclusion

The main aim of the thesis was to confirm hypothesis that Chinese stock markets underperformed during the past. This was examined on two time samples of different length. The first one being longer (1996 – 2013) and second one shorter (2003 – 2013). Unfortunately it was not possible to go further into the past due to lack of data. Still, the dataset was wide enough and therefore I was reasonable to assume it contains some information.

As for the first hypothesis, the prediction of stock index prices was done based on the estimated models. While using FE models with all the explanatory variables resulted in extremely varying results with huge standard errors, BMA method has helped to improve the models estimation dramatically. Subsequently, results obtained from the reduced models worked much more reliably on both lengths of time sample. As for the hypothesis, underperformance was proven for SSE by all models. In case of SZSE, the results differed mainly because of quite good performance of SZSE stock index. However, as described in Chapter 3, SZSE was considerably smaller than SSE throughout whole analyzed period (in fact it was just a fraction of SSE in the beginning). A theoretical weighted index would be much more influenced by SSE and therefore underperformance would be reliably confirmed. CSI 300 was introduced later, so it was not possible to predict its development on a reasonably long sample.

Concerning individual determinants, the third hypothesis was not confirmed as P/E ratio was rarely of any significance (measuring by PIP). In fact it was never picked to be included in a top models given by BMA. Making inference based on the FE models containing all the variables does not seem reasonable.

This research was largely motivated by GDP growth vs. stock markets growth. After the analysis, we can state there is no significant impact of the former on the latter. In this respect, the thesis is in line with most of the previous research mentioned in literature review.

On the other hand, nearly all models yielded inflation and growth of reserves as major determinants of stock market development. As the analysis was working with the nominal index prices, influence of inflation was highly anticipated. Its negative coefficient given by some models, however, was not. As for reserves, it was an unexpected determinant that proved to be very significant. This is one of the key finding of the thesis.

On this simple model, it was confirmed that Chinese stock markets underperformed during the analyzed period. However, it is necessary to be aware of limitation imposed by the model. First of all, it may be subject to omitted variable bias which is very likely scenario when estimating such a complex model. It is also important to mention that underperformance proven by this model does not necessarily mean undervaluation (and therefore a signal to buy Chinese stocks). It can rather mean that Chinese stocks are fairly valued but there are some country-specific factors that were not captured by the model (Sohn et al., 2012). These can include factors such as corporate governance, regulatory environment, self-dealing, reporting unreliability and many more.

Bibliography

ALLEN, F., QIAN, J., SHAN, S. & ZHU, J. Explaining the Underperformance of the Chinese Stock Market. Working paper, 2014

BALTAGI, B. Econometric Analysis of Panel Data. 3rd edition. John Wiley and Sons Ltd, 2001

BREUSCH T. & PAGAN A. The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics. Review of Economic Studies, 1980

CHOW, G.C. China's Economic Transformation. Blackwell, 2007

CHINA SECURITIES INDEX COMPANY. 2014. Available at http://www.csindex.com.cn/sseportal_en/csiportal/indexquery.do

CLARK, T.S. & LINZER, D.A. Should I Use Fixed or Random Effects. Emory University, 2012

DIMSON, E., MARSH, P. & STAUNTON, M. Triumph of the Optimists: 101 Years of Global Investment Returns. eBook, 2002. ISBN: 9780691091945

FENG-CHENG, F., CHU-PING, C. V., YONG-SHENG, C. Productivity and efficiency of state-owned enterprises in China. Peking University, 2007

FERNANDEZ, C., LEY, E., & STEEL, M. F. Benchmark Priors for Bayesian Model Averaging. Journal of Econometrics, 2001

GREENE, W.H. Econometric Analysis, 7th edition, Prentice Hall, 2012

HOETING, J. A., MADIGAN, D., RAFTERY, A. E., & VOLINSKY, C. T. Bayesian Model Averaging: A Tutorial. Statistical Science, 1999

INTERNATIONAL MONETARY FUND. World Economic Outlook: Uneven Growth Short- and Long-Term Factors. IMF, 2015

KOLLER, T., M. GOEDHART, T. E. COPELAND, D. WESSELS and McKinsey and Company. Valuation: Measuring and Managing the Value of Companies. John Wiley & Sons Inc., Hoboken, N.J., 2005

LEY, E. & STEEL, M. F. On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. Journal of Applied Econometrics, 2009

O'NEILL, J., STUPNYTSKA, A. & WRISDALE, J. Linking GDP Growth and Equity Returns. Goldman Sachs Asset Management Monthly Insights, 2011

SANDTE, H. Stock Markets vs. GDP Growth: A Complicated Mixture. BNY Mellon Asset Management, 2012

SOHN, C., TSUI, A., ZHANG, F. & ZHANG, Z., 2012. An Empirical Assessment of A-Share IPO Under-Pricing in China, Seoul Journal of Business, Volume 18, Number 1

TRANSPARENCY INTERNATIONAL. Corruption Perception Index. TI, 2015

THE WORLD BANK. The World Bank Data. WB, 2015. Available at http://data.worldbank.org/

THE WORLD BANK. World Government Indicators. WB, 2014

TSAY, R.S. Analysis of Financial Time Series. Wiley, 2002

WADE, K. & MAY, A. GDP growth and equity market returns. Schroders, 2013

WOOLDRIDGE, J. Econometric Analysis of Cross-Section and Panel Data (2nd edition). Boston: MIT Press, 2010

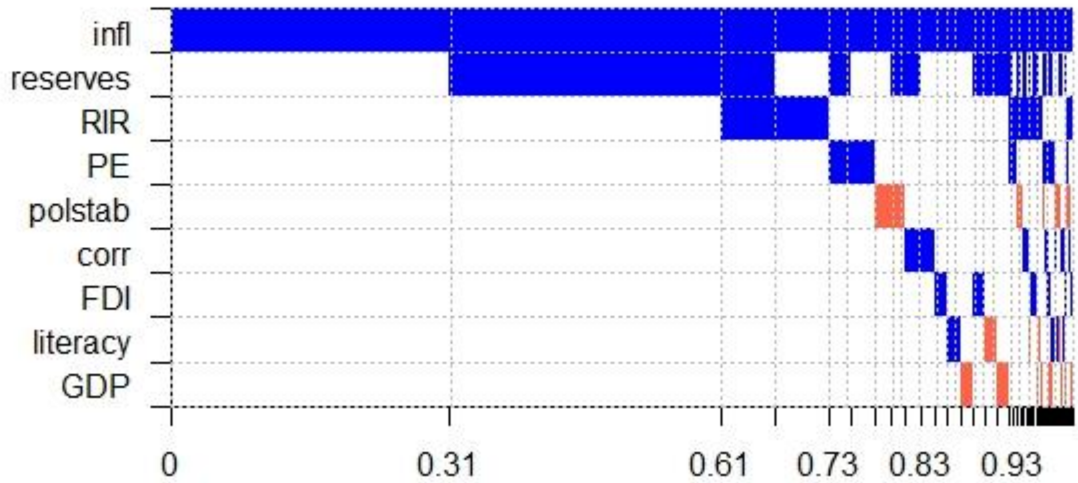
WORLD FEDERATION OF EXCHANGES. World Federation of Exchanges, 2015. Available at http://www.world-exchanges.org/statistics/.

Appendix A

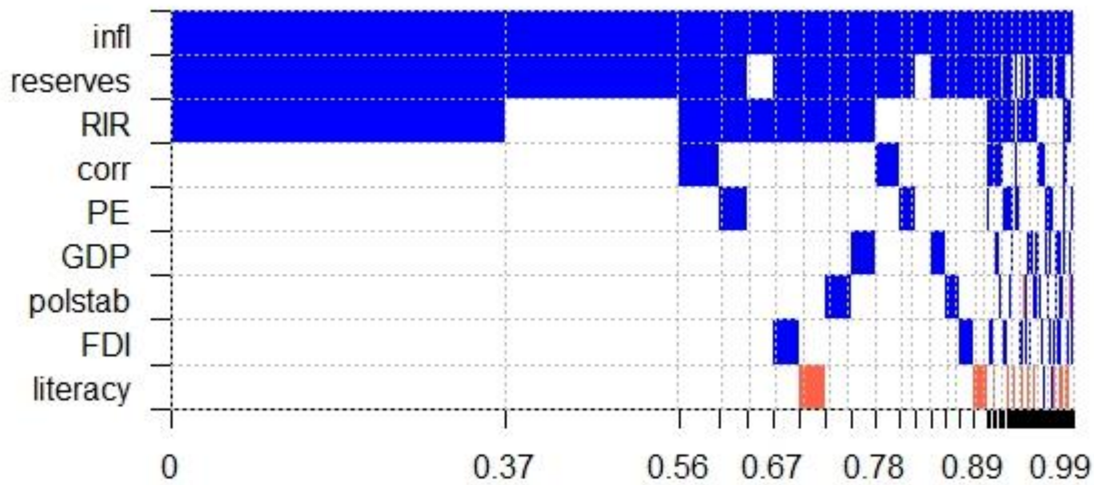
Figures

Figure A.1.: Cumulative model probabilities (1996 – 2013)

Full sample model



Asian countries subsample model



Big economies subsample model

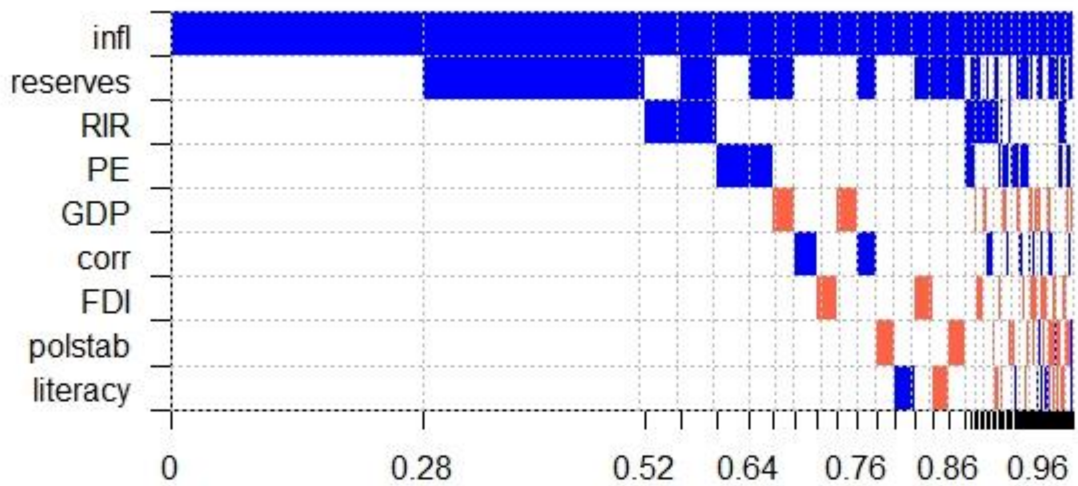
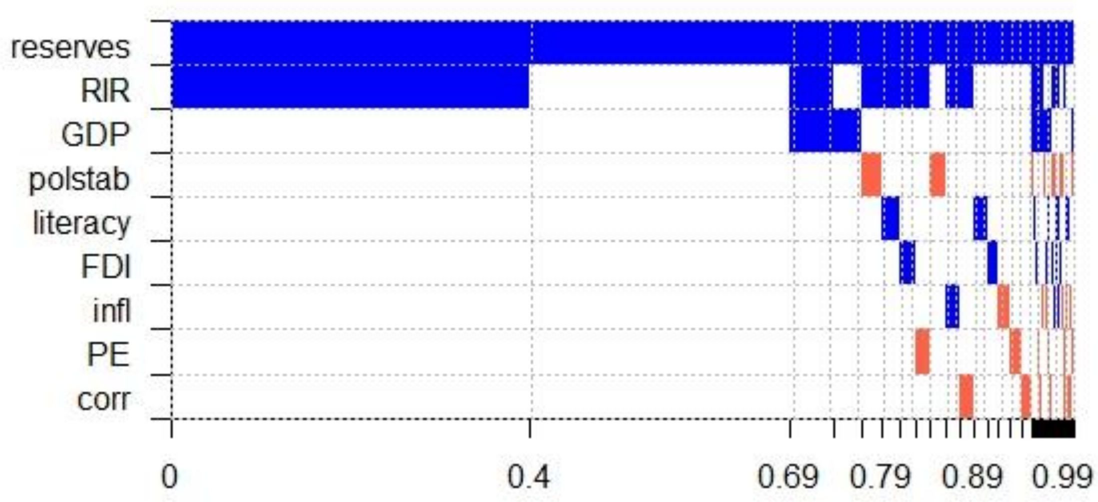
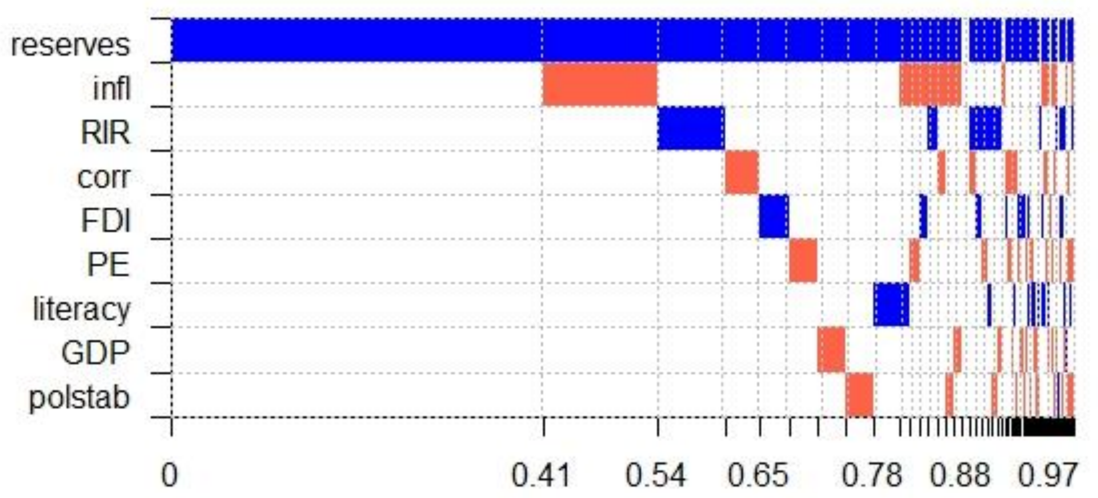


Figure A.2.: Cumulative model probabilities (2003 – 2013)

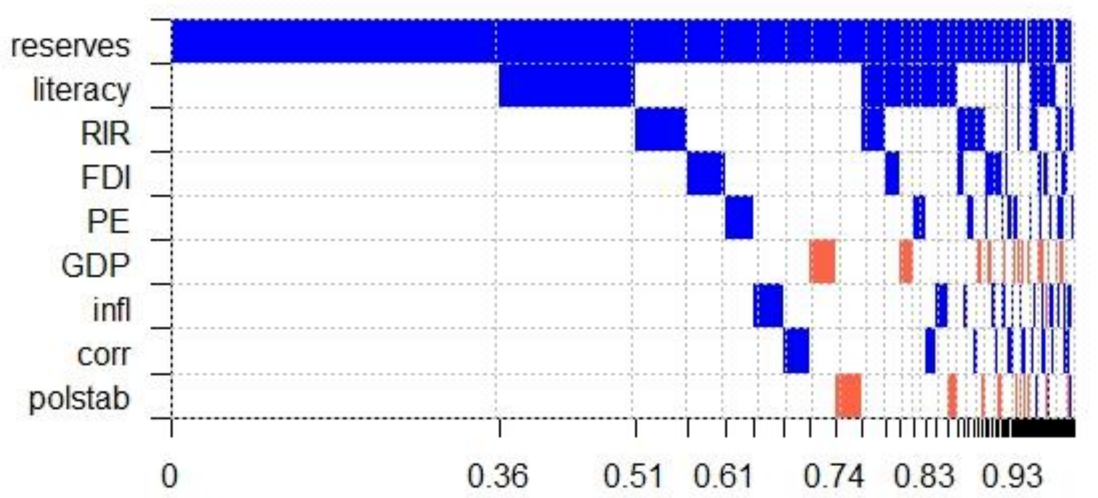
Full sample model



Asian countries subsample model



Big economies subsample model



BRIC countries model

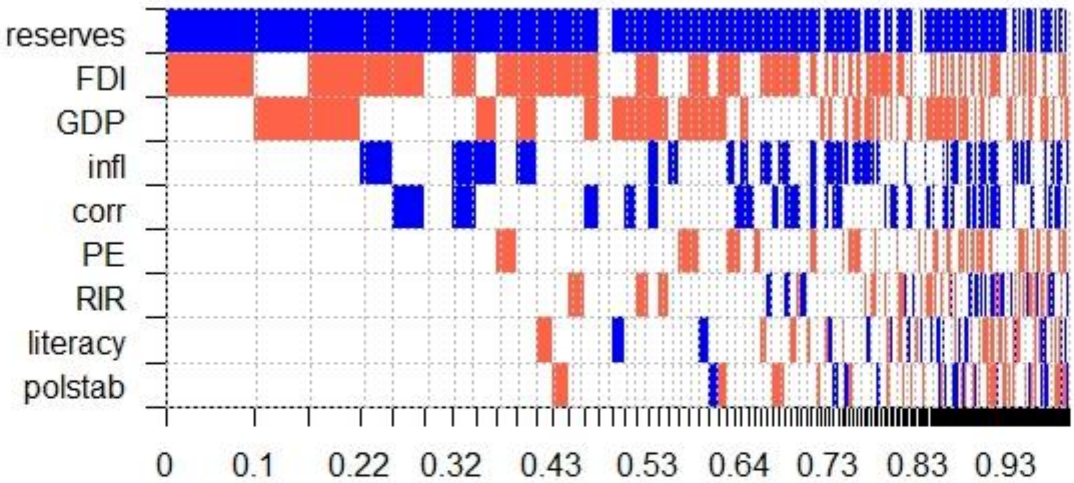
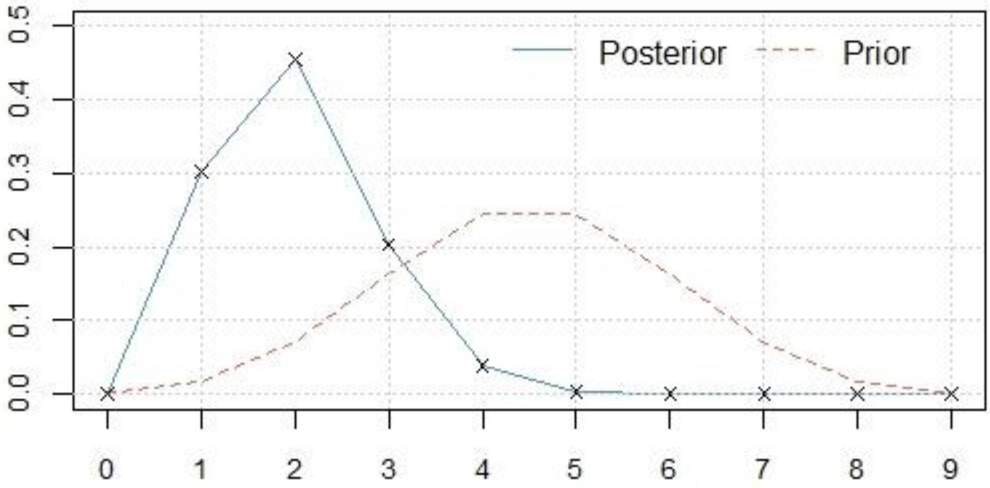
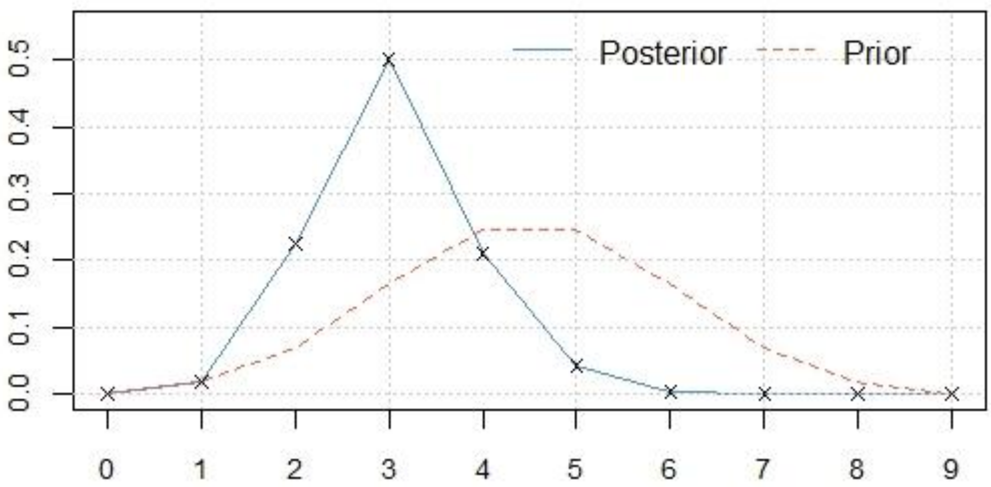


Figure A.3.: Posterior model size distribution (1996 – 2013)

Full sample model



Asian countries subsample model



Big economies subsample model

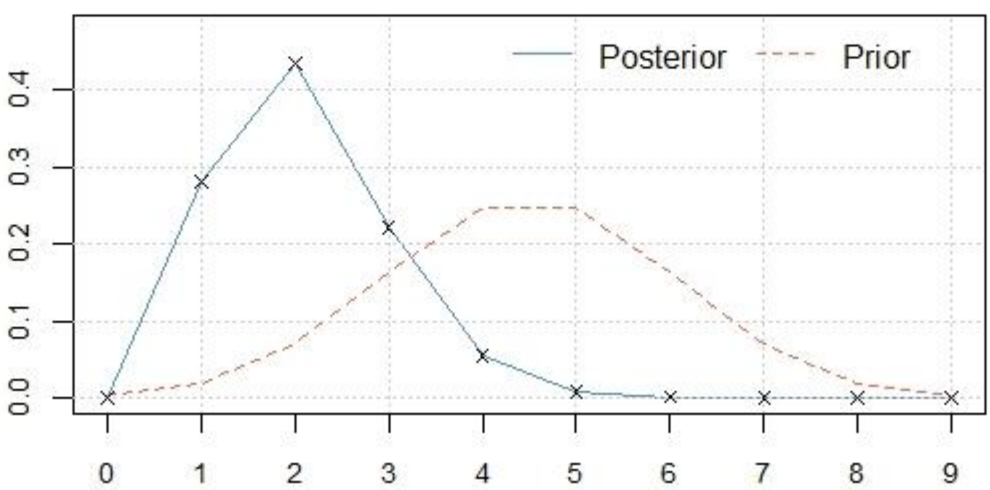
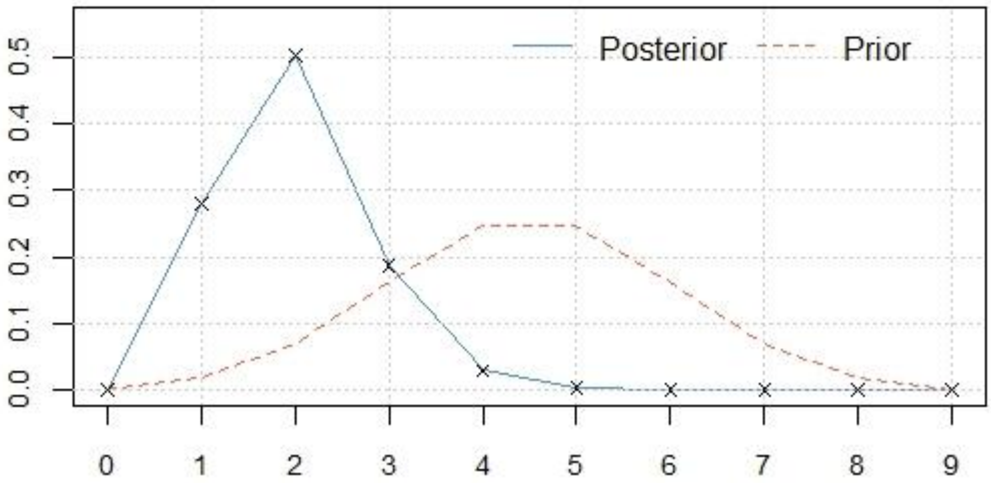
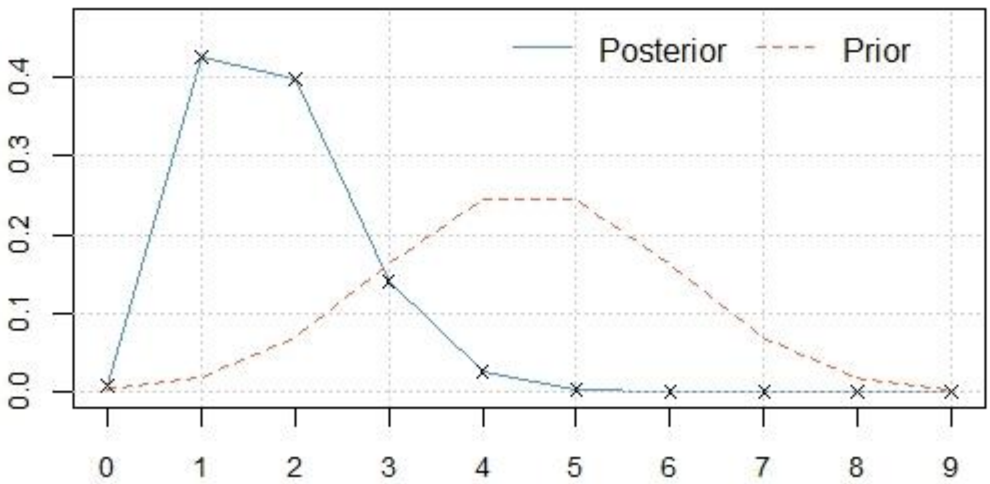


Figure A.4.: Posterior model size distribution (2003 – 2013)

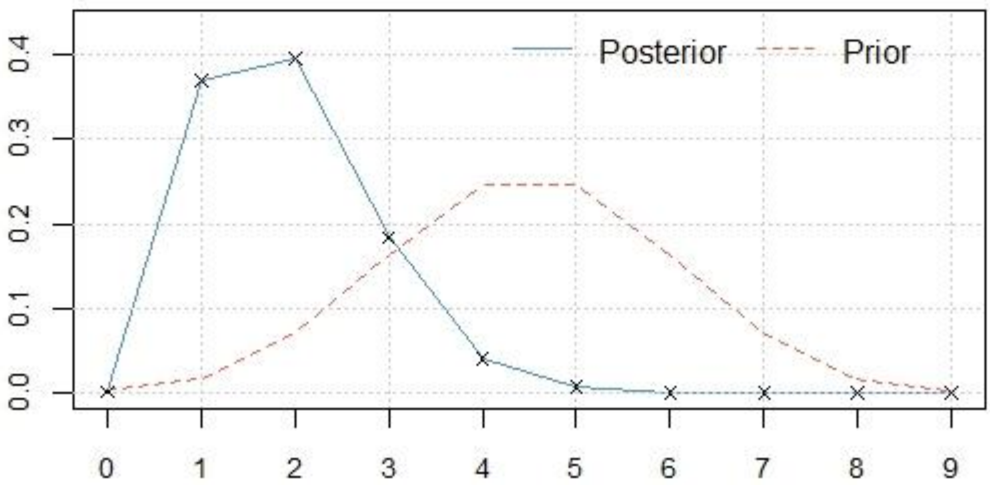
Full sample model



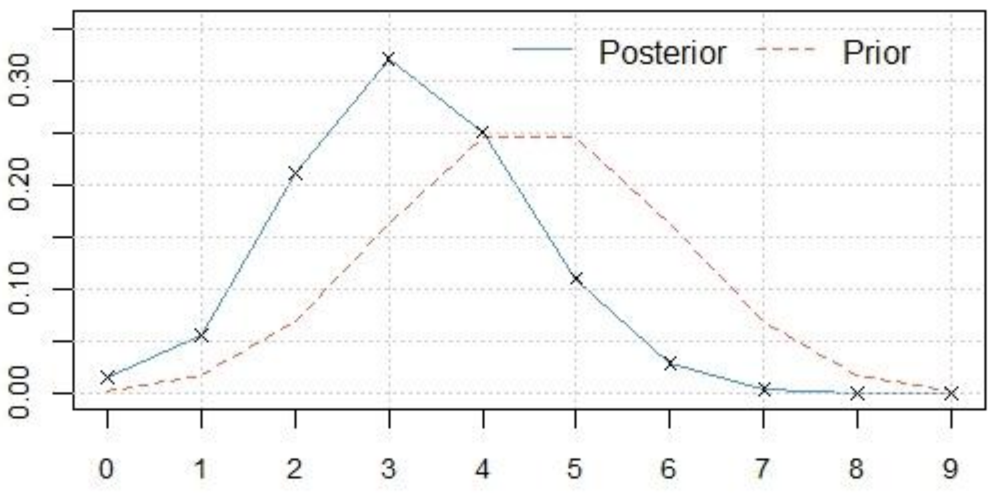
Asian countries subsample model



Big economies subsample model



BRIC countries subsample model



Appendix B
Tables

Table B.1.: Full list major countries‘ stock indices

COUNTRY	INDEX	TICKER
Argentina	Buenos Aires SE Merval Index	.MERV
Australia	FTSE All Share Industrials Index	.FTASX2000
Bahrain	Bahrain All Share Index	.BAX
Belgium	BEL 20 Index	.BFX
Botswana	Botswana Stock Exchange FC Index	.FCIBT
Brazil	Sao Paulo SE Bovespa Index	.BVSP
Bulgaria	Bulgarian Stock Exchange SOFIX Index	.SOFIX
Canada	S&P/TSX 60 INDEX	.SPTSE
Colombia	Colombia SE General Index	.IGBC
Croatia	CROBEX Index	.CRBEX
Denmark	OMX Copenhagen_PI	.OMXCPI
Egypt	EGX 30 Index	.EGX30
Estonia	OMX Tallinn_GI	.OMXTGI
Finland	Finland Index	.dMIFI00000GUS
France	CAC 40 Index	.FCHI
Germany	Deutsche Boerse DAX Index	.GDAXI
Greece	ASE Composite Total Return Index	.RETM
Hungary	Budapest SE Index	.BUX
Chile	IGPA Index	.IGPA
India	S&P BSE Sensex Index	.BSESN
Indonesia	Jakarta SE Composite Index	.JKSE
Ireland	ISEQ Overall Price Index	.ISEQ
Israel	Tel Aviv 25 Index	.TA25
Italy	FTSE MIB Index	.FTMIB
Japan	Nikkei 225 Index	.N225
Jordan	Amman Stock Exchange All-Share Index	.AMMAN
Kenya	Nairobi SE 20 Share Index	.NSE20
Korea	Korea SE Kospi 200 Index	.KS200
Kuwait	Kuwait Stock Exchange (KSX) 15 Index	.KW15
Latvia	OMX Riga_GI	.OMXRGI
Lebanon	S&P Lebanon BMI Banks Index	.SPFLBD4010
Malaysia	FTSE Bursa Malaysia KLCI Index	.KLSE
Mauritius	Semdex Index	.MDEX
Mexico	IPC Index	.MXX
Morocco	Casablanca SE All Share Index	.MASI
Netherlands	Amsterdam Exchanges Index	.AEX
Nigeria	NSE 30 Index	.NGSE30
Norway	Oslo Stock Exchange Equity Index	.OBX
New Zealand	New Zealand Se Top50 Free Index	.NZ50
Oman	Muscat SE General Index	.MSI
Pakistan	Karachi SE 100 Index	.KSE
Peru	Lima SE General Index	.IGRA

Philippines	Philippine SE Composite Index	.PSI
Poland	Warsaw SE WIG Poland Index	.WIG
Portugal	Euronext Lisbon PSI 20 Index	.PSI20
Qatar	Qatar Exchange General Index	.QSI
Romania	Bucharest SE BET Index	.BETI
Russia	MICEX 10 Index	.MCX10
Saudi Arabia	Tadawul FF Index	.TASI
Serbia	Belex15 Belgrade Index	.BELEX15
Singapore	FTSE Straits Times Index	.FTSTI
Slovakia	SAX Index	.SAX
Slovenia	Ljubljana Stock Exchange SBI TOP Index	.SBITOP
South Africa	FTSE/JSE Household Goods & Home Construction sector Index	.JHOUS
Spain	IBEX 35 Index	.IBEX
Sweden	OMX Stockholm 60_GI Index	.OMXS60
Switzerland	Swiss Market Index	.SSMI
Taiwan	FTSE Twse Taiwan 50 Index	.TSE50
Tanzania	Tanzania Share Index	.TSI
Thailand	SET Index	.SETI
Tunisia	Tunis Stock Exchange Weighted Capitalisation Index	.TUNINDEX
Turkey	BIST 100 Index	.XU100
United Kingdom	FTSE 100 Index	.FTSE
Ukraine	PFTS Index	.PFTSI
United States	Dow Jones Industrial Average Index	.DJI
Vietnam	Vietnam Index	.VNI